

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (Other than US) AMRAD OPERATIONS PTY LTD  
(US Only)

(ii) TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS

(iii) NUMBER OF SEQUENCES: 49

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE  
(B) STREET: 1 LITTLE COLLINS STREET  
(C) CITY: MELBOURNE  
(D) STATE: VICTORIA  
(E) COUNTRY: AUSTRALIA  
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT INTERNATIONAL  
(B) FILING DATE: 31-OCT-1997

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO5117  
(B) FILING DATE: 14-FEB-1997

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO 3384  
(B) FILING DATE: 01-NOV-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES DR, E JOHN L  
(C) REFERENCE/DOCKET NUMBER: EJH/EK

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777  
(B) TELEFAX: +61 3 9254 2770

00962560-103197

- 108 -

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGCCGCCC ACGTGAAGGC

20

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTCGCCAATG ACAAGACGCT

20

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1236 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..636

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGAGGCTCAA	GCTCCGGGCG	GATTCTGCGT	GCCGCTCTCG	CTCCTTGGGG	TCTGTTGGCC	-101
GGCCTGTGCC	ACCCGGACGC	CCGGCTCACT	GCCTCTGTCT	CCCCCATCAG	CGCAGCCCCG	-41
GACGCTATGG	CCCACCCCTC	CAGCTGGCCC	CTCGAGTAGG			-1
ATG GTA GCA	CGC AAC CAG	GTG GCA GCC	GAC AAT GCG	ATC TCC CCG	GCA	48
Met Val Ala	Arg Asn Gln	Val Ala Ala	Asp Asn Ala	Ile Ser Pro	Ala	
1	5	10	15			
GCA GAG CCC	CGA CGG CGG	TCA GAG CCC	TCC TCG TCC	TCT TCG TCC		96
Ala Glu Pro	Arg Arg Arg	Ser Glu Pro	Ser Ser Ser	Ser Ser Ser		
	20	25	30			
TCG CCA GCG	GCC CCC GTG	CGT CCC CGG	CCC TGC CCG	GCG GTC CCA	GCC	144

25 FEB 1998 09:52:56

P:\OPER\NEWSOC\1 PRV - 31/10/97

- 109 -

Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala  
 35 40 45  
 CCA GCC CCT GGC GAC ACT CAC TTC CGC ACC TTC CGC TCC CAC TCC GAT 192  
 Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp  
 50 55 60  
 TAC CGG CGC ATC ACG CGG ACC AGC GCG CTC CTG GAC GCC TGC GGC TTC 240  
 Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe  
 65 70 75 80  
 TAT TGG GGA CCC CTG AGC GTG CAC GGG GCG CAC GAG CGG CTG CGT GCC 288  
 Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala  
 85 90 95  
 GAG CCC GTG GGC ACC TTC TTG GTG CGC GAC AGT CGT CAA CGG AAC TGC 336  
 Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys  
 100 105 110  
 TTC TTC GCG CTC AGC GTG AAG ATG GCT TCG GGC CCC ACG AGC ATC CGC 384  
 Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg  
 115 120 125  
 GTG CAC TTC CAG GCC GGC CGC TTC CAC TTG GAC GGC AGC CGC GAG ACC 432  
 Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr  
 130 135 140  
 TTC GAC TGC CTT TTC GAG CTG CTG GAG CAC TAC GTG GCG GCG CCG CGC 480  
 Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg  
 145 150 155 160  
 CGC ATG TTG GGG GCC CCG CTG CGC CAG CGC CGC GTG CGG CCG CTG CAG 528  
 Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln  
 165 170 175  
 GAG CTG TGT CCG CAG CGC ATC GTG GCC GCC GTG GGT CGC GAG AAC CTG 576  
 Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu  
 180 185 190  
 GCG CGC ATC CGT CTT AAC CCG GTA CTC CGT GAC TAC CTG AGT TCC TTC 624  
 Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe  
 195 200 205  
 CCC TTC CAG ATC TGA CCGGCTG CCGCTGTGCC GCAGCATTA GTGGGGGCGC 676  
 Pro Phe Gln Ile \*  
 210  
 CTTATTATTT CTTATTATTA ATTATTATTA TTTTCTGGA ACCACGTGGG AGCCCTCCCC 736  
 GCCTGGGTCG GAGGGAGTGG TTGTGGAGGG TGAGATGCCT CCCACTTCTG GCTGGAGACC 796  
 TCATCCACC TCTCAGGGGT GGGGGTGCTC CCCTCCTGGT GCTCCCTCCG GGTCCCCCCT 856  
 GGTGTAGCA GCTTGTGTCT GGGGCCAGGA CCTGAATTCC ACTCCTACCT CTCCATGTTT 916  
 ACATATTCCC AGTATCTTTG CACAAACCAG GGGTCGGGGA GGGTCTCTGG CTTCAATTTT 976  
 CTGCTGTGCA GAATATCCTA TTTTATATTT TTACAGCCAG TTTAGGTAAT AAACCTTATT 1036  
 ATGAAAGTTT TTTTITAAAA GAAAAAAAAA AAAAAAAAAA 1075

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 212 amino acids

08962550-10319

- 110 -

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala  
1 5 10 15  
Ala Glu Pro Arg Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser Ser  
20 25 30  
Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala  
35 40 45  
Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp  
50 55 60  
Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe  
65 70 75 80  
Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala  
85 90 95  
Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys  
100 105 110  
Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg  
115 120 125  
Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr  
130 135 140  
Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg  
145 150 155 160  
Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln  
165 170 175  
Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu  
180 185 190  
Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe  
195 200 205  
Pro Phe Gln Ile  
210

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1121 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 223..819

08962550-103197

[illegible]

- 112 -

AAGATGTAGC TAGGTATTTT AAAGTTCCCC TTAGGTAGTT TTAGCTGAAT GATGCTTTCT 1046  
 TTCCTATGGC TGCTCAAGAT CAAATGGCCC TTTTAAATGA AACAAAACAA AACAAAACAA 1106  
 AAAAAAAAAA AAAAAA 1121

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 198 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Thr Leu Arg Cys Leu Glu Pro Ser Gly Asn Gly Ala Asp Arg Thr  
 1 5 10 15  
 Arg Ser Gln Trp Gly Thr Ala Gly Leu Pro Glu Glu Gln Ser Pro Glu  
 20 25 30  
 Ala Ala Arg Leu Ala Lys Ala Leu Arg Glu Leu Ser Gln Thr Gly Trp  
 35 40 45  
 Tyr Trp Gly Ser Met Thr Val Asn Glu Ala Lys Glu Lys Leu Lys Glu  
 50 55 60  
 Ala Pro Glu Gly Thr Phe Leu Ile Arg Asp Ser Ser His Ser Asp Tyr  
 65 70 75 80  
 Leu Leu Thr Ile Ser Val Lys Thr Ser Ala Gly Pro Thr Asn Leu Arg  
 85 90 95  
 Ile Glu Tyr Gln Asp Gly Lys Phe Arg Leu Asp Ser Ile Ile Cys Val  
 100 105 110  
 Lys Ser Lys Leu Lys Gln Phe Asp Ser Val Val His Leu Ile Asp Tyr  
 115 120 125  
 Tyr Val Gln Met Cys Lys Asp Lys Arg Thr Gly Pro Glu Ala Pro Arg  
 130 135 140  
 Asn Gly Thr Val His Leu Tyr Leu Thr Lys Pro Leu Tyr Thr Ser Ala  
 145 150 155 160  
 Pro Thr Leu Gln His Phe Cys Arg Leu Ala Ile Asn Lys Cys Thr Gly  
 165 170 175  
 Thr Ile Trp Gly Leu Pro Leu Pro Thr Arg Leu Lys Asp Tyr Leu Glu  
 180 185 190  
 Glu Tyr Lys Phe Gln Val  
 195

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2187 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

0092560-1039

- 113 -

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 18..695

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCTGGCTCC	GTGCGCC	ATG	GTC	ACC	CAC	AGC	AAG	TTT	CCC	GCC	GCC	GGG	50
		Met	Val	Thr	His	Ser	Lys	Phe	Pro	Ala	Ala	Gly	
		1				5					10		
ATG AGC CGC	CCC CTG GAC	ACC	AGC	CTG	CGC	CTC	AAG	ACC	TTC	AGC	TCC	98	
Met Ser Arg	Pro Leu Asp	Thr	Ser	Leu	Arg	Leu	Lys	Thr	Phe	Ser	Ser		
	15			20					25				
AAA AGC GAG	TAC CAG CTG	GTG	GTG	AAC	GCC	GTG	CGC	AAG	CTG	CAG	GAG	146	
Lys Ser Glu	Tyr Gln Leu	Val	Val	Asn	Ala	Val	Arg	Lys	Leu	Gln	Glu		
	30		35					40					
AGC GGA TTC	TAC TGG AGC	GCC	GTG	ACC	GGC	GGC	GAG	GCG	AAC	CTG	CTG	194	
Ser Gly Phe	Tyr Trp Ser	Ala	Val	Thr	Gly	Gly	Glu	Ala	Asn	Leu	Leu		
	45		50				55						
CTC AGC GCC	GAG CCC GCG	GGC	ACC	TTT	CTT	ATC	CGC	GAC	AGC	TCG	GAC	242	
Leu Ser Ala	Glu Pro Ala	Gly	Thr	Phe	Leu	Ile	Arg	Asp	Ser	Ser	Asp		
	60		65			70					75		
CAG CGC CAC	TTC TTC ACG	TTG	AGC	GTC	AAG	ACC	CAG	TCG	GGG	ACC	AAG	290	
Gln Arg His	Phe Phe Thr	Leu	Ser	Val	Lys	Thr	Gln	Ser	Gly	Thr	Lys		
	80				85				90				
AAC CTA CGC	ATC CAG TGT	GAG	GGG	GGC	AGC	TTT	TCG	CTG	CAG	AGT	GAC	338	
Asn Leu Arg	Ile Gln Cys	Glu	Gly	Gly	Ser	Phe	Ser	Leu	Gln	Ser	Asp		
	95			100					105				
CCC CGA AGC	ACG CAG CCA	GTT	CCC	CGC	TTC	GAC	TGT	GTA	CTC	AAG	CTG	386	
Pro Arg Ser	Thr Gln Pro	Val	Pro	Arg	Phe	Asp	Cys	Val	Leu	Lys	Leu		
	110		115					120					
GTG CAC CAC	TAC ATG CCG	CCT	CCA	GGG	ACC	CCC	TCC	TTT	TCT	TTG	CCA	434	
Val His His	Tyr Met Pro	Pro	Pro	Gly	Thr	Pro	Ser	Phe	Ser	Leu	Pro		
	125		130				135						
CCC ACG GAA	CCC TCG TCC	GAA	GTT	CCG	GAG	CAG	CCA	CCT	GCC	CAG	GCA	482	
Pro Thr Glu	Pro Ser Ser	Glu	Val	Pro	Glu	Gln	Pro	Pro	Ala	Gln	Ala		
	140		145			150					155		
CTC CCC GGG	AGT ACC CCC	AAG	AGA	GCT	TAC	TAC	ATC	TAT	TCT	GGG	GGC	530	
Leu Pro Gly	Ser Thr Pro	Lys	Arg	Ala	Tyr	Tyr	Ile	Tyr	Ser	Gly	Gly		
	160			165						170			
GAG AAG ATT	CCG CTG GTA	CTG	AGC	CGA	CCT	CTC	TCC	TCC	AAC	GTG	GCC	578	
Glu Lys Ile	Pro Leu Val	Leu	Ser	Arg	Pro	Leu	Ser	Ser	Asn	Val	Ala		
	175			180					185				
ACC CTC CAG	CAT CTT TGT	CGG	AAG	ACT	GTC	AAC	GGC	CAC	CTG	GAC	TCC	626	
Thr Leu Gln	His Leu Cys	Arg	Lys	Thr	Val	Asn	Gly	His	Leu	Asp	Ser		
	190		195				200						
TAT GAG AAA	GTG ACC CAG	CTG	CCT	GGA	CCC	ATT	CGG	GAG	TTC	CTG	GAT	674	
Tyr Glu Lys	Val Thr Gln	Leu	Pro	Gly	Pro	Ile	Arg	Glu	Phe	Leu	Asp		
	205		210			215							
CAG TAT GAT	GCT CCA CTT	TAAGGAGCAA	AAGGGTCAGA	GGGGGGCCTG								722	
Gln Tyr Asp	Ala Pro Leu												
	220		225										

00962560 "10319"

GGTCGGTCCG TCGCCTCTCC TCCGAGGCAC ATGGCACAAG CACAAAAATC CAGCCCCAAC 782  
GGTCGGTAGC TCCCAGTGAG CCAGGGGCAG ATTGGCTTCT TCCTCAGGCC CTCCACTCCC 842  
GCAGAGTAGA GCTGGCAGGA CCTGGAATTC GTCTGAGGGG AGGGGGAGCT GCCACCTGCT 902  
TTCCCCCTC CCCCAGCTCC AGCTTCTTTC AAGTGAGGCC AGCCGGCCTG GCCTGGTGGG 962  
ACAATACCTT TGACAAGCGG ACTCTCCCTT CCCCTTCCTC CACACCCCTT CTGCTTCCCA 1022  
AGGGAGGTGG GSACACCTCC AAGTGTGAA CTTAGAACTG CAAGGGGAAT CTTCAAACCTT 1082  
TCCCCTGGA ACTTGTGTTGC GCTTTGATTT GGTGTTGATCA AGAGCAGGCA CCTGGGGGAA 1142  
GGATGGAAGA GAAAAGGGTG TGTGAAGGGT TTTTATGCTG GCCAAAGAAA TAACCACTCC 1202  
CACTGCCCAA CCTAGGTGAG GAGTGGTGGC TCCTGGCTCT GGGGAGAGTG GCAAGGGGTG 1262  
ACCTGAAGAG AGCTATACTG GTGCCAGGCT CCTCTCCATG GGGCAGCTAA TGAAACCTCG 1322  
CAGATCCCTT GCACCCCAAGA ACCCTCCCCG TTGTGAAGAG GCAGTAGCAT TTAGAAGGGA 1382  
GACAGATGAG GCTGGTGAGC TGGCCGCCTT TTCCAACACC GAAGGGAGGC AGATCAACAG 1442  
ATGAGCCATC TTGGAGCCCA GGTTCCTCCT GGAGCAGATG GAGGGTTCTG CTTTGTCTCT 1502  
CCTATGTGGG GCTAGGAGAC TCGCCTTAAA TGCCCTCTGT CCCAGGGATG GGGATTGGCA 1562  
CACAAAGGAG CAAACACAGC CAATAGGCAG AGAGTTGAGG GATTCAACCA GGTGGCTACA 1622  
GGCCAGGGGA AGTGGCTGCA GGGGAGAGAC CCAGTCACTC CAGGAGACTC CTGAGTTAAC 1682  
ACTGGAAGA CATTGGCCAG TCCTAGTCAT CTCTCGGTCA GTAGGTCCGA GAGCTTCCAG 1742  
GCCCTGCACA GCCCTCCTTT CTCACCTGGG GGGAGGCAGG AGGTGATGGA GAAGCCTTCC 1802  
CATGCCGCTC ACAGGGGCCT CACGGGAATG CAGCAGCCAT GCAATTACCT GGAACCTGGT 1862  
CTGTCTTGGG GAGAAACAAG TTTTCTGAAG TCAGGTATGG GGCTGGGTGG GGCAGCTGTG 1922  
TGTGGGGTG GCTTTTCTCT CTCTGTTTGA AATAATGTTT ACAATTTGCC TCAATCACTT 1982  
TTATAAAAT CCACCTCCAG CCCGCCCTC TCCCCACTCA GGCCTTCGAG GCTGTCTGAA 2042  
GATGCTTGAA AAATCAACC AAATCCCAGT TCAACTCAGA CTTTGCACAT ATATTTATAT 2102  
TTATACTCAG AAAAGAAACA TTTAGTAAT TTATAATAAA AGAGCACTAT TTTTAAATGA 2162  
AAAAAAAAA AAAAAAAAAA AAAAA 2187

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 225 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Thr His Ser Lys Phe Pro Ala Ala Gly Met Ser Arg Pro Leu  
1 5 10 15  
Asp Thr Ser Leu Arg Ile Lys Thr Phe Ser Ser Lys Ser Glu Tyr Gln  
20 25 30

08962560-10319

91



- 115 -

Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu Ser Gly Phe Tyr Trp  
35 40 45  
Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu Leu Ser Ala Glu Pro  
50 55 60  
Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp Gln Arg His Phe Phe  
65 70 75 80  
Thr Leu Ser Val Lys Thr Gln Ser Gly Thr Lys Asn Leu Arg Ile Gln  
85 90 95  
Cys Glu Gly Gly Ser Phe Ser Leu Gln Ser Asp Pro Arg Ser Thr Gln  
100 105 110  
Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu Val His His Tyr Met  
115 120 125  
Pro Pro Pro Gly Thr Pro Ser Phe Ser Leu Pro Pro Thr Glu Pro Ser  
130 135 140  
Ser Glu Val Pro Glu Gln Pro Pro Ala Gln Ala Leu Pro Gly Ser Thr  
145 150 155 160  
Pro Lys Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly Glu Lys Ile Pro Leu  
165 170 175  
Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala Thr Leu Gln His Leu  
180 185 190  
Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser Tyr Glu Lys Val Thr  
195 200 205  
Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp Gln Tyr Asp Ala Pro  
210 215 220  
Leu  
225

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1094 base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCCGGCTGG	CCCCTTCTGT	AGGATGGTAG	CACACAACCA	GGTGGCAGCC	GACAATGCAG	60
TCTCCACAGC	AGCAGAGCCC	CGACGGCGGC	CAGAACCTTC	CTCCTCTTCC	TCCTCCTCGC	120
CCGCGGCCCC	CGCGCGCCCC	CGGCCGTGCC	CCGCGGTCCC	GGCCCCGGCC	CCCGGCGACA	180
CGCACTTCCG	CACATTCCGT	TCGCACGCCG	ATTACCGGCG	CATCACGCGC	GCCAGCGCGC	240
TCCTGGACGC	CTGCGGATTC	TACTGGGGGC	CCCTGAGCGT	GCACGGGGCG	CACGAGCGGC	300
TGCGCGCCGA	GCCCGTGGGC	ACCTTCCTGG	TGCGCGACAG	CCGCCAGCGG	AACTGCTTTT	360
TCGCCCTTAG	CGTGAAGATG	GCCTCGGGAC	CCACGAGCAT	CCGCGTGCAC	TTTCAGGCCG	420
GCCGCTTTCA	CCTGGATGGC	AGCCGCGAGA	GCTTCGACTG	CCTCTTCGAG	CTGCTGGAGC	480

08962560 103197

ACTACGTGGC GGGGCCGCGC CGCATGCTGG GGGCCCCGCT GCGCCAGCGC CGCGTGCGGC	540
CGCTGCAGGA GCTGTGCCGC CAGCGCATCG TGGCCACCGT GGGCCGCGAG AACCTGGCTC	600
GCATCCCCCT CAACCCCGTC CTCCGCGACT ACCTGAGCTC CTTCCCCTTC CAGATTTGAC	660
CGGCAGCGCC CGCCGTGCAC GCAGCATTAA CTGGGATGCC GTGTTATTTT GTTATTACTT	720
GCCTGGAACC ATGTGGGTAC CCTCCCCGCG CTGGGTTGGA GGGAGCGGAT GGGTGTAGGG	780
GCGAGGCGCC TCCCCCCTC GGCTGGAGAC GAGGCCGCGAG ACCCCTTCTC ACCTCTTGAG	840
GGGGTCTCTC CCCTCCTGGT GCTCCCTCTG GGTCCCCCTG GTTGTTGTAG CAGCTTAACT	900
GTATCTGGAG CCAGGACCTG AACTCGCACC TCCTACCTCT TCATGTTTAC ATATACCCAG	960
TATCTTTGCA CAAACCAGGG GTTGGGGGAG GGTCTCTGGC TTTATTTTTC TGCTGTGCAG	1020
AATCCTATTT TATATTTTTT AAAGTCAGTT TAGGTAATAA ACTTTATTAT GAAAGTTTTT	1080
TTTTTTAAAA AAAA	1094

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Val Ala His Asn Gln Val Ala Ala Asp Asn Ala Val Ser Thr Ala  
1 5 10 15

Ala Glu Pro Arg Arg Arg Pro Glu Pro Ser Ser Ser Ser Ser Ser Ser  
20 25 30

Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Ala Val Pro Ala Pro  
35 40 45

Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ala Asp Tyr  
50 55 60

Arg Arg Ile Thr Arg Ala Ser Ala Leu Leu Asp Ala Cys Gly Phe Tyr  
65 70 75 80

Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala Glu  
85 90 95

08962560-103197

91

- 117 -

Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys Phe  
100 105 110

Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg Val  
115 120 125

His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Ser Phe  
130 135 140

Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg Arg  
145 150 155 160

Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln Glu  
165 170 175

Leu Cys Arg Gln Arg Ile Val Ala Thr Val Gly Arg Glu Asn Leu Ala  
180 185 190

Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe Pro  
195 200 205

Phe Gln Ile  
210

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2807 base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAAACCGAG	GCGGGGAGAC	CAGGAGGCCT	TGGCCTCAGA	GCTTCAGAGT	CGCGTGGCAG	60
CAAACAGAGA	AACCTGTAGA	GGCAGTGTG	CGTCACTTAG	CTCAGGGAAG	CTGCACGCGA	120
AACTCACCCG	CCTTCATTCA	TAAACATCGT	CAGCTAGGCA	CCTACTCCTG	GGCTTTCAGG	180
ACAAACTGAA	TCACGAAACC	ACAGTGTCTT	TAAATAGGT	CTGACCGCCT	GAATCCCTGG	240
CCAAGGTGTG	TACGGGGCAT	GGGAGCCCTT	GTGCAGAGAT	GCTTGCAGGA	GCCTTGAGGG	300
GCTCTGTAAG	ACAGAGGCTA	GGAAGACAAA	GTTGGGGGCT	ACAGCTTCTT	GTCTTGCCCCG	360
GGGCCTCAGT	TTCTTCGGTT	GCCCACGTAG	GAGTGCAGAG	AGTCCAGCCC	CTGGGGACCC	420
AACCCAACCC	CGCCAGTTT	CCGAGGAACT	CGTCCGGGAG	CGGGGGCGCC	CCTCCCGCAC	480
CGCCTTAGGC	TTCTTTTGA	GCCTCTGCGG	TCAGGCCACC	GCTTCCTGGG	AAGCCCAAGC	540

00962560.103197

- 118 -

CAAGGCCAGG	CCGAGTGGCC	AACGGGAGGG	GCCCCGCCGC	GATTCTGGAG	GAGGGCGGCG	600
GCCCCACAGG	TCTCCAGGGC	TGGCTAGCCG	GGCTCCTAGA	GCGGAGACTG	CCAAGGCCCTT	660
CGGGTCTTGG	GCAGGAAGCA	TCCTGGCAGG	GAGGAGTTGC	TTGGGGGGTG	GGGGGGAAAG	720
GCTCCAGGCG	CGGTGGAGCT	CTGACCAGGA	GAATGCACAC	ACTCGGAGGG	GAGGAGGCCGT	780
GTCAGCCCCA	AGCTAGCATC	CCACCCGGGG	AGCAGCGATG	TGGGGCGAAG	GTAGCCAGAG	840
CAAAAGAGCA	GGCACCAGGT	GACACGAAAC	AGAAGATTCC	GGGTAGAGCC	AGAACCCCCAG	900
AAGTCCCATT	CAGGGAAGGT	GCGAGGCGAG	AACGAGTTAG	GTGGACCCCTC	TCCAGGGGCA	960
GCCAAAGAAA	TCTAAAGAGA	ACCCGAAGGA	CTTGCCCGAA	AGAGAAACCG	AAAGCGGCGG	1020
TGGGCGGGAT	CGGTGGGCGG	GGCCTCCCTG	GTTTAAGAGC	TTGATGCAGG	GGCGGGCAGC	1080
AGCAGAGAGA	ACTGCGGCCG	TGGCAGCGGC	ACGGCTCCCCG	GCCCCGGAGC	ATGCGCGACA	1140
GCAGCCCCGG	AACCCCCAGC	CGCGGCGCCC	CGCGTCCCGC	CGCCAGGTGA	GCCGAGGCAG	1200
CTGCGAAGGA	GCAGGCGGGA	GGGGATGGGA	GGAAGGGGAG	CAGAGCCTGG	CAGGACTATC	1260
CTCGCAGACT	GCATGGCGGG	GTCTGTGGATG	CTATGCCTCT	GGCGCCCGCC	CCACCCGGCTG	1320
GCCAGGCGGG	CCCCTCGCGC	GCGCGGGGCG	CCGTCAGCCC	CTCCTCTCCG	GCCCTGAGCC	1380
CGGATCGTCC	GCCCGGGTTC	CAGTTCCCGG	CGTGGCCAGT	AGCGGGCAAC	CGCGAGGCGG	1440
CAAGCCACCC	AGCGGGGACG	GCCTGGAGTC	GGGCCCCCTCT	CCACGCCCCC	TTCTCCACGC	1500
GCGCGGGGAG	GCAGGGCTCC	ACCGCCAGTC	TGGAAGGGTT	CCACATACAG	GAACGGCCTA	1560
CTTCGCAGAT	GAGCCCACCG	AGGCTCAGGC	TCCGGGCGGA	TTCTGCGTGT	CACCCTCGCT	1620
CCTTGGGGTC	CGCTGGCCGG	CCTGTGCCAC	CCGGACGCCC	GCTTCACTGC	CTCTGTCTCC	1680
CCCATCAGCG	CAGCCCCGGA	CGCTATGGCC	CACCCCTCCA	GCTGGCCCCCT	CGAGTAGGAT	1740
GGTAGCACGT	AACCAGGTGG	AAGCCGACAA	TGCGATCTCC	CCGGCATCAG	AGCCCCGACG	1800
GCGGCCAGAG	CCATCCTCGT	CCTCGTCTTC	GTCTCTGCGC	GCGGCCCGGG	CGCGTCCCCG	1860
GCCCTGCCCC	GTGGTCCCGG	CCCCGGCTCC	GGGCGACACT	CACTTCCGCA	CCTTCCGCTC	1920
CCACTCTGAT	TACCGGCGCA	TCACGCGGAC	CAGCGCTCTC	CTGGACGCCT	GCGGCTTCTA	1980
CTGGGGACCC	CTGAGCGTGC	ATGGGGCGCA	CGAACGGCTG	CGTTCCGAAC	CCGTGGGCAC	2040
CTTCTTGGTG	CGCGACAGTC	GCCAGCGGAA	CTGCTTCTTC	GCGCTCAGCG	TGAAGATGGC	2100
TTGGGGCCCC	ACGAGCATTC	GTGTGCACTT	CCAGGCCGGC	CGCTTCCACC	TGGACGGCAA	2160
CCGCGAGACC	TTGACTTGCC	TCTTCGAGCT	GCTGCAACAC	TACGTGGCGG	CGCCGCGCCG	2220
CATGTTGGGG	GCCCCACTGC	GCCAGCGCCG	CGTGCGGGCG	CTGCAGGAGC	TGTGTGCGCA	2280
GCGCATCGTG	GCCGCCGTGG	GTCGCGAGAA	CCTGGCACGC	ATCCCTCTTA	ACCCGGTACT	2340
CCGTGACTAC	CTGAGTTCCCT	TCCCTTTCCA	GATCTGACCG	GCTGCCGCGG	TGCCCCGAGA	2400
ATTAAGTGGG	AGCGCCTTAT	TATTTCTTAT	TATTAATTAT	TATTATTTTT	CTGGAACCAC	2460
GTGGGAGCCC	TCCCCGCCTA	GGTCGGAGGG	AGTGGGTGTG	GAGGGTGAGA	TCCCTCCAC	2520
TTCTGGCTGG	AGACCTTATC	CCGCCTCTCG	GGGGGCCTCC	CCTCCTGGTG	CTCCCTCCCCG	2580
GTCCCCCTGG	TTGTAGCAGC	TTGTGTCTGG	GGCCAGGACC	TGAACTCCAC	GCCTACCTCT	2640
CCATGTTTAC	ATGTTCCAG	TATCTTTGCA	CAAACCAGGG	GTGGGGGAGG	GTCTCTGGCT	2700
TCATTTTTCT	GCTGTGCAGA	ATATTCTATT	TTATATTTTT	ACATCCAGTT	TAGATAATAA	2760
ACTTTATTAT	GAAAGTTTTT	TTTTTTAAAG	AAACAAAGAT	TTCTAGA		2807

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 212 amino acids

03952560.103197

- 119 -

(B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met Val Ala Arg Asn Gln Val Glu Ala Asp Asn Ala Ile Ser Pro Ala
1           5           10           15

Ser Glu Pro Arg Arg Arg Pro Glu Pro Ser Ser Ser Ser Ser Ser Ser
          20           25           30

Ser Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Val Val Pro Ala
          35           40           45

Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp
          50           55           60

Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe
65           70           75           80

Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ser
          85           90           95

Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys
          100          105          110

Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg
          115          120          125

Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Asn Arg Glu Thr
          130          135          140

Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg
145          150          155          160

Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln
          165          170          175

Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu
          180          185          190

```

08962550-103197

- 120 -

Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe  
195 200 205

Pro Phe Gln Ile  
210

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1611 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 263..1529

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGAATTCGG	GCGGGCTGTG	TGAGTCTGTG	AGTGGAAGGC	GCGCCGGCTC	TTTGTCTGA	60
GTGTGACCCG	GTGGCTTTGT	TCCAGGCATT	CCGGTGATTT	CCTCCGGGCA	GTCCGCAGAA	120
---						
GCCGCAGCGG	CCGCCCCGCGC	TCTCTCTGCA	GTCTCCACAC	CCGGGAGAGC	CTGAGCCCGC	180
GTCACGCCCC	TCAGCCCCCG	CTGAGTCCCT	TCTCTGTTGT	CGCGTCCGAA	TCGAGTTCCC	240
GGAATCAGAC	GGTGCCCCAT	AG ATG GCC AGC TTT CCC CCG AGG GTT AAC GAG				292
		Met Ala Ser Phe Pro Pro Arg Val Asn Glu				
		1 5 10				
AAA GAG ATC GTG AGA TCA CGT ACT ATA GGG GAA CTC TTG GCT CCA GCA						340
Lys Glu Ile Val Arg Ser Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala						
	15	20	25			
GCT CCT TTT GAC AAG AAA TGT GGT GGT GAG AAC TGG ACG GTT GCT TTT						388
Ala Pro Phe Asp Lys Lys Cys Gly Gly Glu Asn Trp Thr Val Ala Phe						
	30	35	40			

- 121 -

GCT CCT GAT GGT TCC TAC TTT GCG TGG TCA CAA GGA TAT CGC ATA GTG	436
Ala Pro Asp Gly Ser Tyr Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val	
45 50 55	
AAG CTT GTC CCG TGG TCC CAG TGC CGT AAG AAC TTT CTT TTG CAT GGT	484
Lys Leu Val Pro Trp Ser Gln Cys Arg Lys Asn Phe Leu Leu His Gly	
60 65 70	
TCC AAA AAT GTT ACC AAT TCA AGC TGT CTA AAA TTG GCA AGA CAA AAC	532
Ser Lys Asn Val Thr Asn Ser Ser Cys Leu Lys Leu Ala Arg Gln Asn	
75 80 85 90	
AGT AAT GGT GGT CAG AAA AAC AAG CCT CCT GAG CAC GTT ATA GAC TGT	580
Ser Asn Gly Gly Gln Lys Asn Lys Pro Pro Glu His Val Ile Asp Cys	
95 100 105	
GGA GAC ATA GTC TGG AGT CTT GCT TTT GGG TCT TCA GTT CCA GAA AAA	628
Gly Asp Ile Val Trp Ser Leu Ala Phe Gly Ser Ser Val Pro Glu Lys	
110 115 120	
CAG AGT CGT TGC GTT AAT ATA GAA TGG CAT CGG TTC CGA TTT GGA CAG	676
Gln Ser Arg Cys Val Asn Ile Glu Trp His Arg Phe Arg Phe Gly Gln	
125 130 135	
GAT CAG CTA CTC CTT GCC ACA GGA TTA AAC AAT GGT CGC ATC AAA ATC	724
Asp Gln Leu Leu Leu Ala Thr Gly Leu Asn Asn Gly Arg Ile Lys Ile	
140 145 150	
TGG <del>GAT</del> GTA TAT ACA GGA AAA CTC CTC CTT AAT TTG GTA GAC CAC ATT	772
Trp Asp Val Tyr Thr Gly Lys Leu Leu Leu Asn Leu Val Asp His Ile	
155 160 165 170	
GAA ATG GTT AGA GAT TTA ACT TTT GCT CCA GAT GGG AGC TTA CTC CTT	820
Glu Met Val Arg Asp Leu Thr Phe Ala Pro Asp Gly Ser Leu Leu Leu	
175 180 185	
GTA TCA GCT TCA AGA GAC AAA ACT CTA AGA GTG TGG GAC CTG AAA GAT	868
Val Ser Ala Ser Arg Asp Lys Thr Leu Arg Val Trp Asp Leu Lys Asp	
190 195 200	
GAT GGA AAC ATG GTG AAA GTA TTG CGG GCA CAT CAG AAT TGG GTG TAC	916
Asp Gly Asn Met Val Lys Val Leu Arg Ala His Gln Asn Trp Val Tyr	
205 210 215	
AGT TGT GCA TTC TCT CCC GAC TGT TCT ATG CTG TGT TCA GTG GGC GCC	964

00962550 "10319"

- 122 -

Ser Cys Ala Phe Ser Pro Asp Cys Ser Met Leu Cys Ser Val Gly Ala	
220 225 230	
ACT AAA GCA GTT TTC CTT TGG AAT ATG GAT AAA TAC ACC ATG ATT AGG	1012
Ser Lys Ala Val Phe Leu Trp Asn Met Asp Lys Tyr Thr Met Ile Arg	
235 240 245 250	
AAG CTG GAA GGT CAT CAC CAT GAT GTT GTA GCT TGT GAC TTT TCT CCT	1060
Lys Leu Glu Gly His His His Asp Val Val Ala Cys Asp Phe Ser Pro	
255 260 265	
GAT GGA GCA TTG CTA GCT ACT GCA TCC TAT GAC ACT CGT GTG TAT GTC	1108
Asp Gly Ala Leu Leu Ala Thr Ala Ser Tyr Asp Thr Arg Val Tyr Val	
270 275 280	
TGG GAT CCA CAC AAT GGA GAC CTT CTG ATG GAG TTT GGG CAC CTG TTT	1156
Trp Asp Pro His Asn Gly Asp Leu Leu Met Glu Phe Gly His Leu Phe	
285 290 295	
CCC TCG CCC ACT CCA ATA TTT GCT GGA GGA GCA AAT GAC CGA TGG GTG	1204
Pro Ser Pro Thr Pro Ile Phe Ala Gly Gly Ala Asn Asp Arg Trp Val	
300 305 310	
AGA GCT GTG TCT TTC AGT CAT GAT GGA CTG CAT GTT GCC AGC CTT GCT	1252
Arg Ala Val Ser Phe Ser His Asp Gly Leu His Val Ala Ser Leu Ala	
315 320 325 330	
GAT GAT AAA ATG GTG AGG TTC TGG AGA ATC GAT GAG GAT TGT CCG GTA	1300
Asp Asp Lys Met Val Arg Phe Trp Arg Ile Asp Glu Asp Cys Pro Val	
335 340 345	
CAA GTT GCA CCT TTG AGC AAT GGT CTT TGC TGT GCC TTT TCT ACT GAT	1348
Gln Val Ala Pro Leu Ser Asn Gly Leu Cys Cys Ala Phe Ser Thr Asp	
350 355 360	
GGC AGT GTT TTA GCT GCT GGG ACA CAT GAT GGA AGT GTG TAT TTT TGG	1396
Gly Ser Val Leu Ala Ala Gly Thr His Asp Gly Ser Val Tyr Phe Trp	
365 370 375	
GCC ACT CCA AGG CAA GTC CCT AGC CTT CAA CAT ATA TGT CGC ATG TCA	1444
Ala Thr Pro Arg Gln Val Pro Ser Leu Gln His Ile Cys Arg Met Ser	
380 385 390	
ATC CGA AGA GTG ATG TCC ACC CAA GAA GTC CAA AAA CTG CCT GTT CCT	1492
Ile Arg Arg Val Met Ser Thr Gln Glu Val Gln Lys Leu Pro Val Pro	

08962560 "103197"



- 123 -

395	400	405	410	
TCC AAA ATA TTG GCG TTT CTC TCC TAC CGC GGT TAG A CTGAAGACTG				1539
Ser Lys Ile Leu Ala Phe Leu Ser Tyr Arg Gly *				
	415		420	
CCTTTCCTGG TAGGCCTGCC AGACAGAGCG CCCTTTACAA GACACACCTC AAGCTTTACC				1599
TCGTGCCGAA TT				1611

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Ala	Ser	Phe	Pro	Pro	Arg	Val	Asn	Glu	Lys	Glu	Ile	Val	Arg	Ser	
1				5					10					15		
Arg	Thr	Ile	Gly	Glu	Leu	Leu	Ala	Pro	Ala	Ala	Pro	Phe	Asp	Lys	Lys	
			20					25					30			
Cys	Gly	Gly	Glu	Asn	Trp	Thr	Val	Ala	Phe	Ala	Pro	Asp	Gly	Ser	Tyr	
			35				40					45				
Phe	Ala	Trp	Ser	Gln	Gly	Tyr	Arg	Ile	Val	Lys	Leu	Val	Pro	Trp	Ser	
			50				55				60					
Gln	Cys	Arg	Lys	Asn	Phe	Leu	Leu	His	Gly	Ser	Lys	Asn	Val	Thr	Asn	
			65			70				75				80		
Ser	Ser	Cys	Leu	Lys	Leu	Ala	Arg	Gln	Asn	Ser	Asn	Gly	Gly	Gln	Lys	
				85				90						95		
Asn	Lys	Pro	Pro	Glu	His	Val	Ile	Asp	Cys	Gly	Asp	Ile	Val	Trp	Ser	
				100				105					110			
Leu	Ala	Phe	Gly	Ser	Ser	Val	Pro	Glu	Lys	Gln	Ser	Arg	Cys	Val	Asn	
			115					120					125			

00962560-103197

- 124 -

Ile Glu Trp His Arg Phe Arg Phe Gly Gln Asp Gln Leu Leu Leu Ala  
130 135 140

Thr Gly Leu Asn Asn Gly Arg Ile Lys Ile Trp Asp Val Tyr Thr Gly  
145 150 155 160

Lys Leu Leu Leu Asn Leu Val Asp His Ile Glu Met Val Arg Asp Leu  
165 170 175

Thr Phe Ala Pro Asp Gly Ser Leu Leu Leu Val Ser Ala Ser Arg Asp  
180 185 190

Lys Thr Leu Arg Val Trp Asp Leu Lys Asp Asp Gly Asn Met Val Lys  
195 200 205

Val Leu Arg Ala His Gln Asn Trp Val Tyr Ser Cys Ala Phe Ser Pro  
210 215 220

Asp Cys Ser Met Leu Cys Ser Val Gly Ala Ser Lys Ala Val Phe Leu  
225 230 235 240

Trp Asn Met Asp Lys Tyr Thr Met Ile Arg Lys Leu Glu Gly His His  
245 250 255

His Asp Val Val Ala Cys Asp Phe Ser Pro Asp Gly Ala Leu Leu Ala  
260 265 270

Thr Ala Ser Tyr Asp Thr Arg Val Tyr Val Trp Asp Pro His Asn Gly  
— 275 280 285

Asp Leu Leu Met Glu Phe Gly His Leu Phe Pro Ser Pro Thr Pro Ile  
290 295 300

Phe Ala Gly Gly Ala Asn Asp Arg Trp Val Arg Ala Val Ser Phe Ser  
305 310 315 320

His Asp Gly Leu His Val Ala Ser Leu Ala Asp Asp Lys Met Val Arg  
325 330 335

Phe Trp Arg Ile Asp Glu Asp Cys Pro Val Gln Val Ala Pro Leu Ser  
340 345 350

Asn Gly Leu Cys Cys Ala Phe Ser Thr Asp Gly Ser Val Leu Ala Ala  
355 360 365

08962560 10319

- 125 -

Gly Thr His Asp Gly Ser Val Tyr Phe Trp Ala Thr Pro Arg Gln Val  
 370 375 380

Pro Ser Leu Gln His Ile Cys Arg Met Ser Ile Arg Arg Val Met Ser  
 385 390 395 400

Thr Gln Glu Val Gln Lys Leu Pro Val Pro Ser Lys Ile Leu Ala Phe  
 405 410 415

Leu Ser Tyr Arg Gly \*  
 420

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 783 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTGTCTTCCT CCGCAGCGCG AGGCTGGGTA CAGGGTCTAT TGTCTGTGGT TGA	60
CTTTGGTCTG AGGCCTTCGG GAGCTTTCCC GAGGCAGTTA GCAGAAGCCG CAGCGACCGC	120
CCCCGCCCGT CTCCTCTGTC CCTGGGCCCCG GGAGACAAAC TTGGCGTCAC GCCCTCAGCG	180
GTCGCCACTC TCTTCTCTGT TGTGGGTCC GCATCGTATT CCCGGAATCA GACGGTGCCC	240
CATAGATGGC CAGCTTTCCC CCGAGGGTCA ACGAGAAAGA GATCGTGAGA TCACGTACTA	300
TAGGTGAAC TTTAGCTCCT GCAGCTCCTT TTGACAAGAA ATGTGGTCGT GAAAATTGGA	360
CTGTTGCTTT TGCTCCAGAT GGTTCATACT TTGCTTGGTC ACAAGGACAT CGCACAGTAA	420
AGCTTGTTCC GTGGTCCCAG TGCCTTCAGA ACTTTCTCTT GCATGGCACC AAGAATGTTA	480
CCAATTCAAG CAGTTTAAGA TTGCCAAGAC AAAATAGTGA TGGTGGTCAG AAAAATAAGC	540
CTCGTGACAT ATTATAGACT GTGGAGATAT AGTCTGGAGT CTTGCTTTTG GGTCA	600

08962560 "103197

- 126 -

TCCAGAAAAA CAGAGTCGCT GTGTAAATAT AGAATGGCAT CGCTTCAGAT TTGGACAAGA 660  
 TCAGCTACTT CTTGCTACAG GGTGGAACAA TGGGCGTATC AAAATATGGG ATGTATATCA 720  
 GGAAACTCCT CCTTAACTTG GTAGATCATA CTGAAGTGGT CAGAGATTTA ACTTTTGCTC 780  
 CAG 783

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1122 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTCTGTATGT CTGAATGAAG CTATAACATT TGCCTTTTTA TTGCAGGTTT TCCTTTGGAA 60  
 TATGGATAAA TACACCATGA TACGGAAACT AGAAGGACAT CACCATGATG TGGTAGCTTG 120  
 TGACTTTTCT CCTGATGGAG CATTACTGGC TACTGCATCT TATGATACTC GAGTATATAT 180  
 CTGGGATCCA CATAATGGAG ACATTCTGAT GGAATTTGGG CACCTGTTTC CCCCACCTAC 240  
 TCCAATATTT GCTGGAGGAG CAAATGACCG GTGGGTACGA TCTGTATCTT TTAGCCATGA 300  
 TGGACTGCAT GTTGAAGCC TTGCTGATGA TAAAATGGTG AGGTTCTGGA GAATTGATGA 360  
 GGATTATCCA GTGCAAGTTG CACCTTTGAG CAATGGTCTT TGCTGTGCCT TCTCTACTGA 420  
 TGGCAGTGTT TTAGCTGCTG GGACACATGA CGGAAGTGTG TATTTTGGG CCACTCCACG 480  
 GCAGGTCCTT AGCCTGCAAC ATTTATGTCG CATGTCAATC CGAAGAGTGA TGCCACCCCA 540  
 AGAAGTTCAG GAGCTGCCGA TTCCTTCCAA GCTTTTGGAG TTTCTCTCGT ATCGTATTTA 600  
 GAAGATTCTG CCTTCCCTAG TAGTAGGGAC TGACAGAATA CACTTAACAC AAACCTCAAG 660  
 CTTTACTGAC TTCAATTATC TGTTTTAAA GACGTAGAAG ATTTATTTAA TTTGATATGT 720

08962560-10319

- 127 -

TCTTGACTG CATTGTGATC AGTTGAGCTT TTAATAATATT ATTTATAGAC AATAGAAGTA	780
TTTCTGAACA TATCAAATAT AAATTTTTTTT AAAGATCTAA CTGTGAAAAC ATACATACCT	840
GTACATATTT AGATATAAGC TGCTATATGT TGAATGGACC CTTTTCCTTT TCTGATTTTT	900
AGTTCTGACA TGTATATATT GCTTCAGTAG AGCCACAATA TGTATCTTTG CTGTAAAGTG	960
CAAGGAAATT TTAAATTCTG GGACACTGAG TTAGATGGTA AATACTGACT TACGAAAGTT	1020
GAATTGGGTG AGGCGGGCAA ATCACCTGAG GTCAGCAGTT TGAGACTAGC CTGGCAAACA	1080
TGATGAAACC CTGTCTCTAC TAAAAATACA AAAAAAAAAA AA	1122

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2537 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 422..2029

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGCACGAGC CGGGCTCCGT CCGGAGGAAG CGAGGCTGCG CCGCCGGCCC GGCAGGAGCG	60
GAGGACGGGA GCGCGGGCGG TCGCGCTCGC CCTGTGCTG ACTGCGCTGC CCCGGCCCAT	120
CCTTGCCTGG CCGCAGGTGC CCTGGATGAG GCCGCCGCGC GTGTCCCGGC CGCTGAGTGT	180
CCCCGCGGT CGCCCGGCGC CTGCCCTCAA GCGGCCGCT CTCCTTGCCC GGGTCCCCGT	240
TTTCCCCGG CGCAGTCCTC CTCCGGTGGG CGCCTCCGA CCTCGGCGCA GCGGACACGG	300
CCCTCGGGCC GGGATGGATC CGCCGGGAAG AGGAAGACAA GCCGGGGCGT TGAGCCCCCTG	360

00962560 "103197"

- 128 -

CGCACGGTGC CGCCGCGCGT AGTGGGAGCT TACTCGCAGT AGGCTCTCGC TCTTCTAATC	420
A ATG GAT AAA GTG GGG AAA ATG TGG AAC AAC TTA AAA TAC AGA TGC	466
Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys	
1 5 10 15	
CAG AAT CTC TTC AGC CAC GAG GGA GGA AGC CGT AAT GAG AAC GTG GAG	514
Gln Asn Leu Phe Ser His Glu Gly Gly Ser Arg Asn Glu Asn Val Glu	
20 25 30	
ATG AAC CCC AAC AGA TGT CCG TCT GTC AAA GAG AAA AGC ATC AGT CTG	562
Met Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu	
35 40 45	
GGA GAG GCA GCT CCC CAG CAA GAG AGC AGT CCC TTA AGA GAA AAT GTT	610
Gly Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val	
50 55 60	
GCC TTA CAG CTG GGA CTG AGC CCT TCC AAG ACC TTT TCC AGG CGG AAC	658
Ala Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn	
65 70 75	
CAA AAC TGT GCC GCA GAG ATC CCT CAA GTG GTT GAA ATC AGC ATC GAG	706
Gln Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu	
80 85 90 95	
AAA GAC AGT GAC TCG GGT GCC ACC CCA GGA ACG AGG CTT GCA CGG AGA	754
Lys Asp Ser Asp Ser Gly Ala Thr Pro Gly Thr Arg Leu Ala Arg Arg	
100 105 110	
GAC TCC TAC TCG CGG CAC GCC CCG TGG GGA GGA AAG AAG AAA CAT TCC	802
Asp Ser Tyr Ser Arg His Ala Pro Trp Gly Gly Lys Lys Lys His Ser	
115 120 125	
TGT TCC ACA AAG ACC CAG AGT TCA TTG GAT ACC GAG AAA AAG TTT GGT	850
Cys Ser Thr Lys Thr Gln Ser Ser Leu Asp Thr Glu Lys Lys Phe Gly	
130 135 140	
AGA ACT CGA AGC GGC CTT CAG AGG CGA GAG CGG CGC TAT GGA GTC AGC	898
Arg Thr Arg Ser Gly Leu Gln Arg Arg Glu Arg Arg Tyr Gly Val Ser	
145 150 155	
TCC ATG CAG GAC ATG GAC AGC GTT TCT AGC CGC GCG GTC GGG AGC CGC	946
Ser Met Gln Asp Met Asp Ser Val Ser Ser Arg Ala Val Gly Ser Arg	
160 165 170 175	

09962560-103197

PROPERE/HSCSI.PR.V-31/10/97

- 129 -

TCC CTG AGG CAG AGG CTC CAG GAC ACG GTG GGT TTG TGT TTT CCC ATG	994
Ser Leu Arg Gln Arg Leu Gln Asp Thr Val Gly Leu Cys Phe Pro Met	
180 185 190	
AGA ACT TAC AGC AAG CAG TCA AAG CCA CTC TTT TCC AAT AAA AGA AAA	1042
Arg Thr Tyr Ser Lys Gln Ser Lys Pro Leu Phe Ser Asn Lys Arg Lys	
195 200 205	
ATA CAT CTT TCT GAA TTA ATG CTG GAG AAA TGC CCT TTT CCT GCT GGC	1090
Ile His Leu Ser Glu Leu Met Leu Glu Lys Cys Pro Phe Pro Ala Gly	
210 215 220	
TCG GAT TTA GCA CAA AAG TGG CAT TTG ATT AAA CAG CAT ACC GCC CCT	1138
Ser Asp Leu Ala Gln Lys Trp His Leu Ile Lys Gln His Thr Ala Pro	
225 230 235	
GTG AGC CCA CAC TCA ACA TTT TTT GAT ACA TTT GAT CCA TCA CTG GTG	1186
Val Ser Pro His Ser Thr Phe Phe Asp Thr Phe Asp Pro Ser Leu Val	
240 245 250 255	
TCT ACA GAA GAT GAA GAA GAT AGG CTT CGC GAG AGA AGA CGG CTT AGT	1234
Ser Thr Glu Asp Glu Glu Asp Arg Leu Arg Glu Arg Arg Arg Leu Ser	
260 265 270	
ATC GAA GAA GGG GTG GAT CCC CCT CCC AAC GCA CAA ATA CAC ACC TTT	1282
Ile Glu Glu Gly Val Asp Pro Pro Pro Asn Ala Gln Ile His Thr Phe	
275 280 285	
GAA GGT ACT GCA CAG GTC AAC CCA TTG TAT AAG CTG GGA CCA AAG TTA	1330
Glu Ala Thr Ala Gln Val Asn Pro Leu Tyr Lys Leu Gly Pro Lys Leu	
290 295 300	
GCT CCT GGG ATG ACA GAG ATA AGT GGA GAT GGT TCT GCA ATT CCA CAA	1378
Ala Pro Gly Met Thr Glu Ile Ser Gly Asp Gly Ser Ala Ile Pro Gln	
305 310 315	
GCA ATT GTG ACT CAG AAG AGG ATT CAA CCA CCC TAT GTC TGC AGT CAC	1426
Ala Ile Val Thr Gln Lys Arg Ile Gln Pro Pro Tyr Val Cys Ser His	
320 325 330 335	
GGA GGC AGA AGC AGC GCC AGG TGT CCG GGG ACA GCC ACG CGC ACG TTA	1474
Gly Gly Arg Ser Ser Ala Arg Cys Pro Gly Thr Ala Thr Arg Thr Leu	
340 345 350	
GCA GAC AGG GAG CTT GGA AAG TTC ATA CGC AGA TCG ATT ACA TAC ACT	1522

08962560-103197

- 130 -

Ala Asp Arg Glu Leu Gly Lys Phe Ile Arg Arg Ser Ile Thr Tyr Thr	
355 360 365	
GCC TCG TGC CAG ATT TGC TTC AGA TCA CAG GGA ATC CCT GTT ACT GGG	1570
Ala Ser Cys Gln Ile Cys Phe Arg Ser Gln Gly Ile Pro Val Thr Gly	
370 375 380	
GCG TGA TGG ACC GAT ACG AGG CCG AAG CCC TTC TAG AAG GGA AAC CCG	1618
Ala * Trp Thr Asp Thr Arg Pro Lys Pro Phe * Lys Gly Asn Arg	
385 390 395	
AAG GCA CGT TCT TGC TCA GGG ACT CTG CAC AGG AGG ACT ACC TCT TCT	1666
Lys Ala Arg Ser Cys Ser Gly Thr Leu His Arg Arg Thr Thr Ser Ser	
400 405 410 415	
CTG TGA GCT TCC GCC GCT ACA ACA GGT CTC TGC ACG CCC GGA TCG AGC	1714
Leu * Ala Ser Ala Ala Thr Thr Gly Leu Cys Thr Pro Gly Ser Ser	
420 425 430	
AGT GGA ACC ACA ACT TCA GCT TCG ATG CCC ATG ACC CCT GCG TGT TTC	1762
Ser Gly Thr Thr Thr Ser Ala Ser Met Pro Met Thr Pro Ala Cys Phe	
435 440 445	
ACT CCT CCA CGT CAC GGG GCT TCT CGA ACA CTA TAA AGA CCC CAG CTC	1810
Thr Pro Pro Arg His Gly Ala Ser Arg Thr Leu * Arg Pro Gln Leu	
450 455 460	
TTG CAT GTT TTT TGA ACC GTT GCT AAC GAT ATC ACT GAA TAG AAC TTT	1858
Leu His Val Phe * Thr Val Ala Asn Asp Ile Thr Glu * Asn Phe	
465 470 475	
CCC TTT CAG CCT GCA GTA TAT CTG CCG CGC AGT GAT CTG CAG ATG CAC	1906
Pro Phe Gln Pro Ala Val Tyr Leu Pro Arg Ser Asp Leu Gln Met His	
480 485 490 495	
TAC GTA TGA TGG GAT TGA CGG GCT CCC GCT ACC GTC GAT GTT ACA GGA	1954
Tyr Val * Trp Asp * Arg Ala Pro Ala Thr Val Asp Val Thr Gly	
500 505 510	
TTT TTT AAA AGA GTA TCA TTA TAA ACA AAA AGT TAG GGT TCG CTG GTT	2002
Phe Phe Lys Arg Val Ser Leu * Thr Lys Ser * Gly Ser Leu Val	
515 520 525	
AGA ACG AGA CCA GTC AAA GCA AAG TAACTCCTGT CCCCAAAGGG CACTAACTAA	2056
Arg Thr Arg Pro Val Lys Ala Lys	

08962560-103197



- 131 -

530

535

GTCTGCTCCT CCCGTGCATC GAACTGCACC CATAGGAGGC AGTCAGCTGC TAGGATTTCC 2116  
 CACCCAGAAT GGGAGCTTAG TCATTAGCCT CTGCCCTATG GGGTCCGCTG TTCCTCAGAC 2176  
 AAAGGTGCCT AGGGACAGCA AGATGGCTTG CAGGTGTTTC GTGGGCTGTG ACAACTGAGG 2236  
 GAGGCAACTC TGGGGCATTG GCTATGAAGA ATTCTATTTT TTACCGAAGA ACAAATTATT 2296  
 AATATTGGAT GGGTATTTCA ATAGTGTGAC TAATGTTTGA AATTATTTTT TCTAAGAATT 2356  
 TTTCTATAAC CTTCAGAAAA AGTAGTGATG TTTGTAGTTA CTATAAATCA AGCTTTGAAA 2416  
 GTTCAAAACA AACAAGTTAA ATAAAAGACT ACCTTCCTTT TAGAGAAAAC AAATGCAAGT 2476  
 TTTCCAGCC ACAGGCATTG TGAAGTGTTA ATGTTGCTTG TTATCAGCTC CTTTCTCCTC 2536  
 C 2537

(2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 535 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys Gln  
 1 5 10 15  
 Asn Leu Phe Ser His Glu Gly Gly Ser Arg Asn Glu Asn Val Glu Met  
 20 25 30  
 Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu Gly  
 35 40 45  
 Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val Ala  
 50 55 60  
 Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn Gln

08962560-103197

- 132 -

65		70		75		80
Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu Lys						
	85		90		95	
Asp Ser Asp Ser Gly Ala Thr Pro Gly Thr Arg Leu Ala Arg Arg Asp						
	100		105		110	
Ser Tyr Ser Arg His Ala Pro Trp Gly Gly Lys Lys Lys His Ser Cys						
	115		120		125	
Ser Thr Lys Thr Gln Ser Ser Leu Asp Thr Glu Lys Lys Phe Gly Arg						
	130		135		140	
Thr Arg Ser Gly Leu Gln Arg Arg Glu Arg Arg Tyr Gly Val Ser Ser						
	145		150		155	160
Met Gln Asp Met Asp Ser Val Ser Ser Arg Ala Val Gly Ser Arg Ser						
		165		170		175
Leu Arg Gln Arg Leu Gln Asp Thr Val Gly Leu Cys Phe Pro Met Arg						
	180		185		190	
Thr Tyr Ser Lys Gln Ser Lys Pro Leu Phe Ser Asn Lys Arg Lys Ile						
	195		200		205	
His Leu Ser Glu Leu Met Leu Glu Lys Cys Pro Phe Pro Ala Gly Ser						
	210		215		220	
Asp Leu Ala Gln Lys Trp His Leu Ile Lys Gln His Thr Ala Pro Val						
	225		230		235	240
Ser Pro His Ser Thr Phe Phe Asp Thr Phe Asp Pro Ser Leu Val Ser						
	245		250		255	
Thr Glu Asp Glu Glu Asp Arg Leu Arg Glu Arg Arg Arg Leu Ser Ile						
	260		265		270	
Glu Glu Gly Val Asp Pro Pro Pro Asn Ala Gln Ile His Thr Phe Glu						
	275		280		285	
Ala Thr Ala Gln Val Asn Pro Leu Tyr Lys Leu Gly Pro Lys Leu Ala						
	290		295		300	
Pro Gly Met Thr Glu Ile Ser Gly Asp Gly Ser Ala Ile Pro Gln Ala						

00962560-103497

- 133 -

305                      310                      315                      320  
 Ile Val Thr Gln Lys Arg Ile Gln Pro Pro Tyr Val Cys Ser His Gly  
                          325                      330                      335  
 Gly Arg Ser Ser Ala Arg Cys Pro Gly Thr Ala Thr Arg Thr Leu Ala  
                          340                      345                      350  
 Asp Arg Glu Leu Gly Lys Phe Ile Arg Arg Ser Ile Thr Tyr Thr Ala  
                          355                      360                      365  
 Ser Cys Gln Ile Cys Phe Arg Ser Gln Gly Ile Pro Val Thr Gly Ala  
                          370                      375                      380  
 \* Trp Thr Asp Thr Arg Pro Lys Pro Phe \* Lys Gly Asn Arg Lys  
 385                      390                      395                      400  
 Ala Arg Ser Cys Ser Gly Thr Leu His Arg Arg Thr Thr Ser Ser Leu  
                          405                      410                      415  
 \* Ala Ser Ala Ala Thr Thr Gly Leu Cys Thr Pro Gly Ser Ser Ser  
                          420                      425                      430  
 Gly Thr Thr Thr Ser Ala Ser Met Pro Met Thr Pro Ala Cys Phe Thr  
                          435                      440                      445  
 Pro Pro Arg His Gly Ala Ser Arg Thr Leu \* Arg Pro Gln Leu Leu  
                          450                      455                      460  
 His Val Phe \* Thr Val Ala Asn Asp Ile Thr Glu \* Asn Phe Pro  
 465                      470                      475                      480  
 Phe Gln Pro Ala Val Tyr Leu Pro Arg Ser Asp Leu Gln Met His Tyr  
                          485                      490                      495  
 Val \* Trp Asp \* Arg Ala Pro Ala Thr Val Asp Val Thr Gly Phe  
                          500                      505                      510  
 Phe Lys Arg Val Ser Leu \* Thr Lys Ser \* Gly Ser Leu Val Arg  
                          515                      520                      525  
 Thr Arg Pro Val Lys Ala Lys  
                          530                      535

00962560-103197

- 134 -

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1221 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATTAAACAG CATACAGCTC CTGTGAGCCC ACATTCAACA TTTTGTGATA CTTTGATCCA	60
TCTTTGGTTT CTACAGAAGA TGAAGAAGAT AGGCTTAGAG AGAGAAGGCG GCTTAGTATT	120
GAAGAAGGGG TTGATCCCCC TCCAATGCA CAAATACATA CATTTGAAGC TACTGCACAG	180
GTTAATCCAT TATTAACTG GGACCAAAAT TAGCTCCTGG AATGACTGAA ATAAGTGGGG	240
ACAGTTCTGC AATTCCACAA GCTAATTGTG ACTCGGAAGA GGATACAACC ACCCTGTGTT	300
GCAGTCACGG AGGCAGAAGC AGCGTCAGAT ATCTGGAGAC AGCCATACCC ATGTTAGCAG	360
ACAGGGAGCT TGGAAAGTCC ACACACAGAT TGATTACATA CACTGCTTCG TGCCTGATTT	420
GCTTCAATT ACAGGGAATC CCTGTTACTG GGGACTGATG GACCGTTATG AAGCAGAAGC	480
CCTTCTCGAA GGGAAACCTG AAGGCACGTT TTTGCTCAGG GACTCTGCGC AAGAGGACTA	540
CTTCTTCTCT GTGAGCTTCC GCCGATACAA CAGATCCCTG CATGCCCGAA TTGAGCAGTG	600
GAATCACAAC TTTAGTTTCG ACGCCCATGA CCCGTGTGTA TTCACTCCT CCACTGTAAC	660
GGGACTTTTA GAACATTATA AAGATCCCAG TTCGTGCATG TTTTGTGAAC CATTGCTTAC	720
TATATCACTA AATAGGACTT TCCCTTTTAG CCTGCAGTAT ATCTGTGCGG CGGTAATCTG	780
CAGGTGCACT ACGTATGATG GAATTGATGG GCTCCCTCTA CCCTCAATGT TACAGGATTT	840
TTTAAAAGAG TATCATTATA AACAAAAAGT TAGAGTTCGC TGGTTGGAAC GAGAACCAGT	900
CAAGGCAAAG TAAACTCTCC GGTCCCCAAA GGGTGTTAAC TAGGTCCGCT TTCATGTGCA	960

08962560 "103197"

- 135 -

TCAGACAGTA CACCTATAGC AAGCACACGT AGCAGTGTTA GGCTTTTTCA TACAGTATGT 1020  
 AAGCTTAGTG TTAGTATCTG TCAGATGCTA CCTGCTGTTA CTTATTCAGA TAAACATGGT 1080  
 GCCTATTGGA ACAATAGCGG ATAGAGCTAC AGGTGTTTCA TAAGACTACA AAAACATTTT 1140  
 GCCTATTTTCG CTAACAGTTT GGTTTTAAAT GGCTGTGGTA TTTGAGTGAG GCAACTCTGG 1200  
 GGCATTTGTT ATGAAGAAAT G 1221

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 116..1330

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCACGAGGC GGTGGTGGCG GCGGCGGGCG CGGCCGCGGC GGGGCGGGCG CGGAATGAAG 60  
 GCCCACGGCC CTGGGGGCTG AGGCGCCCGC CGCCTGGGGC GGGCCGCGCG TCCTC ATG 118  
 Met  
 1  
 GAG GCC GGA GAG GAG CCG CTG CTG CTG GCT GAA CTC AAG CCT GGG CGC 166  
 Glu Ala Gly Glu Glu Pro Leu Leu Leu Ala Glu Leu Lys Pro Gly Arg  
 5 10 15  
 CCC CAC CAG TTC GAC TGG AAG TCA AGC TGC GAG ACC TGG AGC GTG GCC 214  
 Pro His Gln Phe Asp Trp Lys Ser Ser Cys Glu Thr Trp Ser Val Ala  
 20 25 30  
 TTC TCG CCA GAC GGT TCC TGG TTC GCC TGG TCT CAA GGA CAC TGC GTG 262

08965560-103197

- 136 -

Phe Ser Pro Asp Gly Ser Trp Phe Ala Trp Ser Gln Gly His Cys Val	
35 40 45	
GTC AAG CTG GTC CCC TGG CCC TTA GAG GAA CAG TTC ATC CCT AAA GGA	310
Val Lys Leu Val Pro Trp Pro Leu Glu Glu Gln Phe Ile Pro Lys Gly	
50 55 60 65	
TTC GAA GCC AAG AGC CGA AGC AGC AAG AAT GAC CCA AAA GGA CGG GGC	358
Phe Glu Ala Lys Ser Arg Ser Ser Lys Asn Asp Pro Lys Gly Arg Gly	
70 75 80	
AGT CTG AAG GAG AAG ACG CTG GAC TGT GGC CAG ATT GTG TGG GGG CTG	406
Ser Leu Lys Glu Lys Thr Leu Asp Cys Gly Gln Ile Val Trp Gly Leu	
85 90 95	
GCC TTC AGC CCG TGG CCC TCT CCA CCC AGC AGG AAA CTC TGG GCA CGT	454
Ala Phe Ser Pro Trp Pro Ser Pro Pro Ser Arg Lys Leu Trp Ala Arg	
100 105 110	
CAC CAT CCC CAG GCG CCT GAT GTT TCT TGC CTG ATC CTG GCC ACA GGT	502
His His Pro Gln Ala Pro Asp Val Ser Cys Leu Ile Leu Ala Thr Gly	
115 120 125	
CTC AAC GAT GGG CAG ATC AAG ATT TGG GAG GTA CAG ACA GGC CTC CTG	550
Leu Asn Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu Leu	
130 135 140 145	
CTT CTG AAT CTT TCT GGC CAC CAA GAC GTC GTG AGA GAT CTG AGC TTC	598
Leu Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser Phe	
150 155 160	
ACG CCC AGC GGC AGT TTG ATT TTG GTC TCT GCA TCC CGG GAT AAG ACA	646
Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys Thr	
165 170 175	
CTT CGA ATT TGG GAC CTG AAT AAA CAC GGT AAG CAG ATC CAG GTG TTA	694
Leu Arg Ile Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln Val Leu	
180 185 190	
TCC GGC CAT CTG CAG TGG GTT TAC TGC TGC TCC ATC TCC CCT GAC TGT	742
Ser Gly His Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser Pro Asp Cys	
195 200 205	
AGC ATG CTG TGC TCT GCA GCT GGG GAG AAG TCG GTC TTT CTG TGG AGC	790
Ser Met Leu Cys Ser Ala Ala Gly Glu Lys Ser Val Phe Leu Trp Ser	

08962560-103197

- 137 -

210	215	220	225	
ATG CGG TCC TAC ACA CTA ATC CGG AAA CTA GAA GGC CAC CAA AGC AGT				838
Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu Gly His Gln Ser Ser				
230		235	240	
GTT GTC TCC TGT GAT TTC TCT CCT GAT TCA GCC TTG CTT GTC ACA GCT				886
Val Val Ser Cys Asp Phe Ser Pro Asp Ser Ala Leu Leu Val Thr Ala				
245		250	255	
TCG TAT GAC ACC AGT GTG ATT ATG TGG GAC CCC TAC ACC GGC GCG AGG				934
Ser Tyr Asp Thr Ser Val Ile Met Trp Asp Pro Tyr Thr Gly Ala Arg				
260		265	270	
CTG AGG TCA CTT CAT CAC ACA CAA CTT GAA CCC ACC ATG GAT GAC AGT				982
Leu Arg Ser Leu His His Thr Gln Leu Glu Pro Thr Met Asp Asp Ser				
275		280	285	
GAC GTC CAC ATG AGC TCC CTG AGG TCC GTG TGC TTC TCA CCT GAA GGC				1030
Asp Val His Met Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu Gly				
290		295	300	305
TTG TAT CTC GCT ACG GTG GCA GAT GAC AGG CTG CTC AGG ATC TGG GCT				1078
Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp Ala				
310		315	320	
CTG GAA CTG AAG GCT CCG GTT GCC TTT GCT CCG ATG ACC AAT GGT CTT				1126
Leu Glu Leu Lys Ala Pro Val Ala Phe Ala Pro Met Thr Asn Gly Leu				
325		330	335	
TGC TGC ACG TTC TTC CCA CAC GGT GGA ATT ATT GCC ACA GGG ACG AGA				1174
Cys Cys Thr Phe Phe Pro His Gly Gly Ile Ile Ala Thr Gly Thr Arg				
340		345	350	
GAT GGC CAT GTC CAG TTC TGG ACA GCT CCC CGG GTC CTG TCC TCA CTG				1222
Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser Leu				
355		360	365	
AAG CAC TTA TGC AGG AAA GCC CTC CGA AGT TTC CTG ACA ACG TAT CAA				1270
Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr Gln				
370		375	380	385
GTC CTA GCA CTG CCA ATC CCC AAG AAG ATG AAA GAG TTC CTC ACA TAC				1318
Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr Tyr				
390		395	400	

08062560-103197

PROPERE/HSOCSI.PRV - 31/10/97

- 138 -

AGG ACT TTC TAGCAGTGCC GGCTCCCCCA CCTCCTGCAG CAGCAGCAGT 1367  
 Arg Thr Phe  
 405

ACAAGGGACT GGCTAGGATG GAGTCAGGCA GCTCACACTG GACCAGTGTG GACCTTCCTT 1427  
 CCTCCCATGG CATGTGCAAG TAGGTCTGCG TGACCCCACT TCTGTGGTGC CGGCCTTACC 1487  
 TCGTCTTCAT CCGTGGTGAG CAGCCTTCGT CAGTCTAGTT GTGTTGAAGC CAAGTGCAGT 1547  
 TGTGGATGTT GCTGGGGTAA TAAAGGCAAG CGGGCTCCAG AGCCTCTCTG GTGGCGGCCA 1607  
 AGCCACACTC CCTTAACTGG GAAGTACCTG CCACGTAGGG CATTTCTGCT GCCTATTTCC 1667  
 AGCCAGCGGC TGCATGGTTT GAAGTTCCTC CGTTGTGGTC AGAAGAACTC TGGTGTTTGG 1727  
 TTCCCTGCTC AGCTGCGCGT GGACTGGGCT GAGCTCCTCA CCATACACTA GTGCCGGCTT 1787  
 TTGTTTCCTG TAAACAGTGG TTGCATGTGT AGAGAAGTAA CAAGCGAGTA TTCAGATCAT 1847  
 ACGAGGAGGC GTTCCTCGGT GCATGACGGT CAGATGGCCA TTTATCAGCA TATTTATTTG 1907  
 TATTTTCTCA GCACATAGTA AGGTACAAC TGTGTTTCTC AATTGTCTCG AAAAAACAGA 1967  
 GTTCTTAAGT GGCCCACTG TGGAGCCAAG TCTAAGTCGT GTGGAGTCAG TGCTGACATC 2027  
 ACTGGCTTGT GCTGTCTGTC ACATGTGTTT GTCTCTGCTG CTTGACCTCA TGGGATGTAC 2087  
 CCTCEAGTTC AACTGCCCAA AACAGACAGC CCCTTCCAAG CACCGTTCTT TGACAGCGGT 2147  
 AGCAGCTACC TATTCAGAC GCCTCACACA AAATCTGCCT TAGAAAGTTA ATATATTTTA 2207  
 AATTATTTTA AAAGAACTC AACATCTTAT TCTTTGGCCT TTCTTAATTG ATGCTTTATG 2267  
 GAGGCAGTGT TAACATTGTA CAGTGTATGC ATAGAGGAGT CTCCTCTATT TGAAGAACAA 2327  
 TGCAAAATGA GGCTTTCATT GAAGGGAAAA AAAAAAAAAA AA 2369

(2) INFORMATION FOR SEQ ID NO:21:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 404 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

00962560.103197



- 139 -

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Glu	Ala	Gly	Glu	Glu	Pro	Leu	Leu	Leu	Ala	Glu	Leu	Lys	Pro	Gly	1	5	10	15
Arg	Pro	His	Gln	Phe	Asp	Trp	Lys	Ser	Ser	Cys	Glu	Thr	Trp	Ser	Val	20	25	30	
Ala	Phe	Ser	Pro	Asp	Gly	Ser	Trp	Phe	Ala	Trp	Ser	Gln	Gly	His	Cys	35	40	45	
Val	Val	Lys	Leu	Val	Pro	Trp	Pro	Leu	Glu	Glu	Gln	Phe	Ile	Pro	Lys	50	55	60	
Gly	Phe	Glu	Ala	Lys	Ser	Arg	Ser	Ser	Lys	Asn	Asp	Pro	Lys	Gly	Arg	65	70	75	80
Gly	Ser	Leu	Lys	Glu	Lys	Thr	Leu	Asp	Cys	Gly	Gln	Ile	Val	Trp	Gly	85	90	95	
Leu	Ala	Phe	Ser	Pro	Trp	Pro	Ser	Pro	Pro	Ser	Arg	Lys	Leu	Trp	Ala	100	105	110	
Arg	His	His	Pro	Gln	Ala	Pro	Asp	Val	Ser	Cys	Leu	Ile	Leu	Ala	Thr	115	120	125	
Gly	Leu	Asn	Asp	Gly	Gln	Ile	Lys	Ile	Trp	Glu	Val	Gln	Thr	Gly	Leu	130	135	140	
Leu	Leu	Leu	Asn	Leu	Ser	Gly	His	Gln	Asp	Val	Val	Arg	Asp	Leu	Ser	145	150	155	160
Phe	Thr	Pro	Ser	Gly	Ser	Leu	Ile	Leu	Val	Ser	Ala	Ser	Arg	Asp	Lys	165	170	175	
Thr	Leu	Arg	Ile	Trp	Asp	Leu	Asn	Lys	His	Gly	Lys	Gln	Ile	Gln	Val	180	185	190	
Leu	Ser	Gly	His	Leu	Gln	Trp	Val	Tyr	Cys	Cys	Ser	Ile	Ser	Pro	Asp	195	200	205	
Cys	Ser	Met	Leu	Cys	Ser	Ala	Ala	Gly	Glu	Lys	Ser	Val	Phe	Leu	Trp	210	215	220	

00962560-103197

- 140 -

Ser Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu Gly His Gln Ser  
225 230 235 240

Ser Val Val Ser Cys Asp Phe Ser Pro Asp Ser Ala Leu Leu Val Thr  
245 250 255

Ala Ser Tyr Asp Thr Ser Val Ile Met Trp Asp Pro Tyr Thr Gly Ala  
260 265 270

Arg Leu Arg Ser Leu His His Thr Gln Leu Glu Pro Thr Met Asp Asp  
275 280 285

Ser Asp Val His Met Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu  
290 295 300

Gly Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp  
305 310 315 320

Ala Leu Glu Leu Lys Ala Pro Val Ala Phe Ala Pro Met Thr Asn Gly  
325 330 335

Leu Cys Cys Thr Phe Phe Pro His Gly Gly Ile Ile Ala Thr Gly Thr  
340 345 350

Arg Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser  
355 360 365

Leu Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr  
370 375 380

Gln Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr  
385 390 395 400

Tyr Arg Thr Phe

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

DB962560.103197

- 141 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GACACTGCAT CGTCAAACTG ATCCCCTGGC CGTTGGAGGA GCAGTTCATC CCTAAAGGGT	60
TTGAAGCCAA AAGCCGAAGT AGCAAAAATG AGACGAAAGG GCGGGGCAGC CCAAAAGAGA	120
AGACGCTGGA CTGTGGTCAG ATTGTCTGGG GGCTGGCCTT CAGCCTGTGC TTTCCCCACC	180
CAGCAGGAAG CTCTGGGCAC GCCACCACCC CCAAGTGCCC GATGTCTCTT GCCTGGTTCT	240
TGCTACGGGA CTCAACGATG GGCAGATCAA GATCTGGGAG GTGCAGACAG GGCTCCTGCT	300
TTTGAATCTT TCCGGCCACC AAGATGTCGT GAGAGATCTG AGCTTCACAC CCAGTGGCAG	360
TTTGATTTTG GTCTCCGCGT CACGGGATAA GACTCTTCGC ATCTGGGACC TGAATAAACA	420
CGGTAAACAG ATTCAAGTGT TATCGGGCCA CCTGCAGTGG GTTTACTGCT GTTCCATCTC	480
CCCAGACTGC AGCATGCTGT GCTCTGCAGC TGCAGAGAAG TCGGTCTTTC TATGGAGCAT	540
GAGGTCTTAC ACGTTAATTC GGAAGCTAGA GGGCCATCAA AGCAGTGTG TCTCTTGTA	600
CTTCTCCCCC GACTCTGCCC TGCTTGTCAC GGCTTCTTAC GATACCAATG TGATTATGTG	660
GGACCTCTAC ACCGGCGAAA GGCTGAGGTC ACTCCACCAC ACCCAGGTTG ACCCCGCCAT	720
GGATGACAGT GACGTCCACA TTAGCTCACT GAGATCTGTG TGCTTCTCTC CAGAAGGCTT	780
GTACCTTGCC ACGGTGGCAG ATGACAGACT CCTCAGGATC TGGGCCCTGG AACTGAAAAC	840
TCCCATTGCA TTTGCTCCTA TGACCAATGG GCTTTGCTGG CACATTTTTT CCACATGGTG	900
GAGTCATTGC CACAGGGACA AGAGATGGCC ACGTCCAGTT CTGGACAGCT CCTAGGGTCC	960
TGTCCTCACT GAAGCACTTA TGCCGGAAG CCCTTCGAAG TTTCCTAACA ACTTACCAAG	1020
TCCTAGCACT GCCAATCCCC AAGAAAATGA AAGAGTTCCT CACATACAGG ACTTTTTAAG	1080
CAACACCACA TCTTGCTCTT CTTTGTAGCA GGGTAAATCG TCCTGTCAAA GGGAGTTGCT	1140
GGAATAATGG GCCAAACATC TGGTCTTGCA TTGAAATAGC ATTTCTTTGG GATTGTGAAT	1200

08962560-103197

- 142 -

AGAATGTAGC AAAACCAGAT TCCAGTGTAC TAGTCATGGA TTTTTC

1246

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACCATGGTTC CAAGTCCTCT CCCCTGTGGT CAAGTTGCCC GAATGTTGGG CCCAAGTGCC	60
TTTTCCTCCT TGGGCTCCC CTCTGACCT GCAGGACAGT TTTCCGGAGC CCATTGCGTA	120
TGAGGTATTA ATTAGCCTTA ACTAAATTAC AGGGGACTCA GAGGCCGTGC TCCTGACCGA	180
TCCAGACACT ATTTTTTTTT TTTTTTTTTA ACAATGGTGT GCATGTGCAG GAAATGACAA	240
ATTTGTATGT CAGATTATAC AAGGATGTAT TCTTAAACCG CATGACTATT CAGATGGCTA	300
CTGAGTTATC AGTGGCCATT TATTAGCATC ATATTTATTT GTATTTTCTC AACAGATGTT	360
AAGGTACAAC TGTGTTTTTC TCGATTATCT AAAAACCATA GTACTTAAAT TGAACAAAAA	420
AA	422

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2019 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

08962550-103197

- 143 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGCAGGAGGC GGGGTCAGGG CGGAGGCTGA GGACCAAGTA GGCATGGCGG AGGGCGGGAC	60
CGGCCCCGAT GGACGGGCGG GCCCGGGACC CGCAGGTCCT AATCTGAAGG AGTGGCTGAG	120
GGAGCAGTTC TGTGACCATC CACTGGAGCA CTGTGACGAT ACAAGACTCC ATGATGCAGC	180
CTATGTAGGG GACCTCCAGA CCCTCAGGAA CCTACTGCAA GAGGAGAGCT ACCGGAGCCG	240
CATCAATGAG AAGTCTGTCT GGTGCTGCGG CTGGCTTCCC TGCACACCAC TGAGGATCGC	300
AGCCACTGCA GGCCATGGGA ACTGTGTGGA CTTCCCTCATA CGCAAAGGGG CCGAGGTGGA	360
CCTGGTGGAT GTCAAGGGGC AGACTGCCCT GTATGTGGCT GTAGTGAACG GGCACCTGGA	420
GAGCACTGAG ATCCTTTTGG AAGCTGGTGC TGATCCCAAC GGCAGCCGGC ACCACCGCAG	480
CACTCCTGTG TACCATGCCT YTCGTGTGGG TAGGGACGAC ATCCTGAAGG CTCTTATCAG	540
GTATGGGGCA GATGTTGATG TCAACCATCA TCTGAATTCT GACACCCGGC CCCCTTTTTC	600
ACGGCGGCTA ACCTCCTTGG TGGTCTGTCC TCTATACATC AGTGCTGCCT ACCATAACCT	660
TCAGTGCTTC AGGCTGCTCT TGCAGGCTGG GGCAAATCCT GACTTCAATT GCAATGGCCC	720
TGTCAACACC CAGGAGTTCT ACAGGGGATC CCCTGGGTGT GTCATGGATG CTGTCCTGCG	780
CCATGGCTGT GAAGCAGCCT TCGTGAGTCT GTTGGTAGAG TTTGGAGCCA ACCTGAACCT	840
GGTGAAGTGG GAATCCCTGG GCCCAGAGGC AAGAGGCAGA AGAAAGATGG ATCCTGAGGC	900
CTTGCAGGTC TTAAAGAGG CCAGAAGTAT TCCCAGGACC TTGCTGAGTT TGTGCCGGGT	960
GGCTGTGAGA AGAGCTCTTG GCAAATACCG ACTGCATCTG GTTCCCTCGC TGCCGCTGCC	1020
AGACCCATA AAGAAGTTT TGCTTTATGA GTAGCATTCAT CATGCAGTGC TGACTGCAAT	1080
GTGGAAGCCG ATCACCTGCA GTGAAACTG ACACAGACTC TGGCATCCTG GGAACCATGG	1140
CCTGTGCTGC CAGCTTGATC CTTGGCTGTC AGTGAAGAAA AAACGGCTGT GTTCTCTTGG	1200
ACTGTGATTC TATCTCAGGT GCTTGGGCCA TCGAACGCTC CTTGAGTCAT TGTCAACTGA	1260

08962560-103197

- 144 -

GAGGCACATA CAAACTTAAT TTTGTTCCCTC TTCAGTCTCT CTGTTTTGGA TTCTTCCTGG	1320
CAATGTGTGC AGCATGGGCT GAGCCTGGTG ATTGCCCTAG TGGGGAAGGC TTTTTTCTCC	1380
AGGCTATGCA TCTATTTATG TTCCTACTTT GCAATTTATT GTTCTTTTAA GGCTTGATAT	1440
CAAAACAGAA AGAGGTTTGT TAAGAAAAGA TATAGGGAGA AAGGAATTCC GGTTCCTGTC	1500
ACTTGCTAGC CTGCTTTCCT TGCCTGGGTT TGTCTGTCTA TGCTGCCTGG TGCACATCCC	1560
TTCTCTTTCG TGCCACTGTT CTATTTTGGG AGTTGTCTTC CGTCTAAGAT GGCTTCTGGG	1620
GTTCTATCTT ATTGCACAGA GGTCCCAGAA CAGTGTTTAT AGGGCACCAT CTGCTCTGCC	1680
AAGGGTTTTTC TGATGTCTTA CCCTGGGGAT CTTCAGACAG TGGTTACCTT TAGGAGACCC	1740
ACCTGGAACCT AACCATTAAAG TGAATGCCCA CATTGAGATC AGGGACCATC TTAATAGTAC	1800
TCACTGCCAG TCCTCACAAG AGAAGATGAC ACGGGTGCTC TCTTCAGACA CTCCCATACA	1860
GGAAGTTGGA AAATGTCTTG GTCACCTGGG TTGTTCCAG GCTACAACCT CTTGGTGTTC	1920
CACTAARACC AGRATATCCT AGTTTTTTGG GTTGACTGTT CCCTCCCCAC TTTCTTGAA	1980
NCCCAATGCC CNTTTGTKTN GGTTGCTTCC CTAAAAKTT	2019

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala	Arg	Gly	Gly	Val	Arg	Ala	Glu	Ala	Glu	Asp	Gln	Val	Gly	Met	Ala
1						5								10	15

00962560-103197

- 145 -

Glu Gly Gly Thr Gly Pro Asp Gly Arg Ala Gly Pro Gly Pro Ala Gly  
20 25 30

Pro Asn Leu Lys Glu Trp Leu Arg Glu Gln Phe Cys Asp His Pro Leu  
35 40 45

Glu His Cys Asp Asp Thr Arg Leu His Asp Ala Ala Tyr Val Gly Asp  
50 55 60

Leu Gln Thr Leu Arg Asn Leu Leu Gln Glu Glu Ser Tyr Arg Ser Arg  
65 70 75 80

Ile Asn Glu Lys Ser Val Trp Cys Cys Gly Trp Leu Pro Cys Thr Pro  
85 90 95

Leu Arg Ile Ala Ala Thr Ala Gly His Gly Asn Cys Val Asp Phe Leu  
100 105 110

Ile Arg Lys Gly Ala Glu Val Asp Leu Val Asp Val Lys Gly Gln Thr  
115 120 125

Ala Leu Tyr Val Ala Val Val Asn Gly His Leu Glu Ser Thr Glu Ile  
130 135 140

Leu Leu Glu Ala Gly Ala Asp Pro Asn Gly Ser Arg His His Arg Ser  
145 150 155 160

Thr Pro Val Tyr His Ala Xaa Arg Val Gly Arg Asp Asp Ile Leu Lys  
— 165 170 175

Ala Leu Ile Arg Tyr Gly Ala Asp Val Asp Val Asn His His Leu Asn  
180 185 190

Ser Asp Thr Arg Pro Pro Phe Ser Arg Arg Leu Thr Ser Leu Val Val  
195 200 205

Cys Pro Leu Tyr Ile Ser Ala Ala Tyr His Asn Leu Gln Cys Phe Arg  
210 215 220

Leu Leu Leu Gln Ala Gly Ala Asn Pro Asp Phe Asn Cys Asn Gly Pro  
225 230 235 240

Val Asn Thr Gln Glu Phe Tyr Arg Gly Ser Pro Gly Cys Val Met Asp  
245 250 255

08962560-103197

- 146 -

Ala Val Leu Arg His Gly Cys Glu Ala Ala Phe Val Ser Leu Leu Val  
260 265 270

Glu Phe Gly Ala Asn Leu Asn Leu Val Lys Trp Glu Ser Leu Gly Pro  
275 280 285

Glu Ala Arg Gly Arg Arg Lys Met Asp Pro Glu Ala Leu Gln Val Phe  
290 295 300

Lys Glu Ala Arg Ser Ile Pro Arg Thr Leu Leu Ser Leu Cys Arg Val  
305 310 315 320

Ala Val Arg Arg Ala Leu Gly Lys Tyr Arg Leu His Leu Val Pro Ser  
325 330 335

Leu Pro Leu Pro Asp Pro Ile Lys Lys Phe Leu Leu Tyr Glu  
340 345 350

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCATCCATGG CGGAGGGCGG CAGCACGACG GCGGGGAGG GCCGGGCTCC GCAGGTCGTA	60
ATCTGAAGGA GTGGCTGAGG GAGCAATTTT GTGATCATCC GCTGGAGCAC TGTGAGGACA	120
CGAGGCTCCA TGATGCAGCT TACGTCGGCG ACCTCCAGAC CCTCAGGAGC CTATTGCAAG	180
AGGAGAGCTA CCGGAGCCGC ATCAACGAGA AGTCTGTCTG GTGCTGTGGC TGGCTCCCCT	240
GCACACCGTT GCGAATCGCG GCCACTGCAG GCCATGGGAG CTGTGTGGAC TTCCTCATCC	300
GGAAGGGGGC CGAGGTGGAT CTGGTGGACG TAAAAGGACA GACGGCCCTG TATGTGGCTG	360

DB962560.103197



- 147 -

TGGTGAACGG GCACCTAGAG AGTACCCAGA TCCTTCTCGA AGCTGGCGCG GACCCCAAC 419

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 595 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAGGAAGAAG AAAAGTGGAC CCTGAGGCCT TGCAGGTCTT TAAAGAGGCC AGAAGTGTTC	60
CCAGAACCTT GCTGTGTCTG TGCCGTGTGG CTGTGAGAAG AGCTCTTGGC AAAACCGGCT	120
TCATCTGATT CCTTCGCTGC CTCTGCCAGA CCCCATAAAG AAGTTTCTAC TCCATGAGTA	180
GACTCCAAGT GCTGCGGTTG ATTCCAGTGA GGCAGAAAGT GATCTGCAGG GAGGTGGACA	240
CCCAGCCCTG AGTGCTGTGC TGCTGCTGGT CTCCTGATGG CTGTTGCTGC AGAAGATGTC	300
CTCGTAGACT GTCATTGCTC CTCAGGTGCC TGGGCCGCTG AACAGTCCTT GGGTCATTGT	360
CAGCTGAGAG GCTTATACTA AAGTTATTAT TGTTTTTCCC AAGTTCTCTG TTCTGGATT	420
TCAGTTGCAT ATTAATGTAA CGGGCCATGG GGTATGTACA TGTAGGGGCT GAGGTTGGAG	480
GCCTACTAAT TTCCTGTAGG GAAGACTCCC AGCACTTCTG GAACTGTGCT TCTCTTTATT	540
TTTCTACTTC TCAATTTGAT GGTTCGATTA AAGCCTTCTA GTATCTCAAT GAAAA	595

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 896 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

08962550-103197

- 148 -

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 4..396

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTG ATG TCC GCA ATT CTG AAG GTT GGA CAC CAC TGC TGG CTG CCT GTG	48
Met Ser Ala Ile Leu Lys Val Gly His His Cys Trp Leu Pro Val	
1 5 10 15	
ACA TCC GCT GTC AAT CCC CAA AGG ATG CTG AGG CCA CCA CCA ACC GCT	96
Thr Ser Ala Val Asn Pro Gln Arg Met Leu Arg Pro Pro Pro Thr Ala	
20 25 30	
GTT TTC AAC TGT GCC GCT TGC TGC TGT CTG TGG GGG CAG ATG CTG ATG	144
Val Phe Asn Cys Ala Ala Cys Cys Cys Leu Trp Gly Gln Met Leu Met	
35 40 45	
AAT ACA TAC CGT GTA GTT CAG CTT CCT GAG GAG GCC AAG GGC TTG GTG	192
Asn Thr Tyr Arg Val Val Gln Leu Pro Glu Glu Ala Lys Gly Leu Val	
50 55 60	
CCA CCA GAG ATT CTA CAG AAG TAC CAT GGA TTC TAC TCT TCC CTC TTT	240
Pro Pro Glu Ile Leu Gln Lys Tyr His Gly Phe Tyr Ser Ser Leu Phe	
-65 70 75	
GCC TTG GTG AGG CAG CCC AGG TCG CTG CAG CAT CTC TGC CGT TGT GCG	288
Ala Leu Val Arg Gln Pro Arg Ser Leu Gln His Leu Cys Arg Cys Ala	
80 85 90 95	
CTC CGC AGT CAC CTG GAG GGC TGT CTG CCC CAT GCA CTA CCG CGC CTT	336
Leu Arg Ser His Leu Glu Gly Cys Leu Pro His Ala Leu Pro Arg Leu	
100 105 110	
CCC CTG CCA CCG CGC ATG CTC CGC TTT CTG CAG CTG GAC TTT GAG GAT	384
Pro Leu Pro Pro Arg Met Leu Arg Phe Leu Gln Leu Asp Phe Glu Asp	
115 120 125	
CTG CTC TAC TAGGCTTGCT GCCCTGTGAA CAAAGCAGAC CCCACCCCCA	433
Leu Leu Tyr	
130	

00962550-103197

- 149 -

CCCCAGGGC ATCTCTCAGC AATGAATGAT GCAAGGCGGT CTGTCTTCAA GTCAGGAGTG 493  
 GACGCCTTGA TCCACACTTG AGAGAAGAGG CCAGATCAGC ACCYGGCTGG TAGTGATNGC 553  
 AGAGGGCACC TGTGCAGATC TGTGTGCGCA CTGGAAATCT CTAGGCTGAA GGCTYAGAGCA 613  
 AATGGTGCAR GTGTTAGTCC TTGGGANGAG AGACAGANGG TGAGAAAGCA AGACAGAGGT 673  
 GAGAGTGCAC ATGTCAAGTG GTAGATTGCC TTAAGAGAAA GCTAAAAAAA GAAAAAGATT 733  
 CGGGCGAACT TCTTTAGGGG TAATGCTGCA GCGTGTTAAA CTGACTGACC AGCGTCCATA 793  
 TCTTTGGACC CTTCCCGGGT GAAAAAGCCC CTTTCATCTC CAGCGCTCCC CAAGGGTGCT 853  
 TAGCAATACC GGGTGCTTTT CTGCCGCAA GTGAGTTACC AAA 896

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Ser Ala Ile Leu Lys Val Gly His His Cys Trp Leu Pro Val Thr  
 1 5 10 15  
 Ser Ala Val Asn Pro Gln Arg Met Leu Arg Pro Pro Pro Thr Ala Val  
 20 25 30  
 Phe Asn Cys Ala Ala Cys Cys Cys Leu Trp Gly Gln Met Leu Met Asn  
 35 40 45  
 Thr Tyr Arg Val Val Gln Leu Pro Glu Glu Ala Lys Gly Leu Val Pro  
 50 55 60  
 Pro Glu Ile Leu Gln Lys Tyr His Gly Phe Tyr Ser Ser Leu Phe Ala  
 65 70 75 80  
 Leu Val Arg Gln Pro Arg Ser Leu Gln His Leu Cys Arg Cys Ala Leu  
 85 90 95

00962560 103197

- 150 -

Arg Ser His Leu Glu Gly Cys Leu Pro His Ala Leu Pro Arg Leu Pro  
                   100                  105                  110

Leu Pro Pro Arg Met Leu Arg Phe Leu Gln Leu Asp Phe Glu Asp Leu  
                   115                  120                  125

Leu Tyr  
           130

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 436 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTGGGGGCGT CATCATGACC TCCTCTAGGG CTCTGCAACA TGACTCCTGT GGTGCAAATC 60

AACAAATTGT TCACTGATGA ATCCACAAGG ATCTCTGGGC CTACAACCAG GTCCTGGTCC 120

ACATGACTGT CGTCTTCGGA GAAGGCACCA CTCGCCCCCG GCAGGTACGG CTGACACCTC 180

CATGGGAGAA GACGTATCCA GGCAGCAGCT GCGCGGCCCT TCAAGAGGGC ACATCCCGTC 240

ATCTAAAGGC ACGGTGTACT GAAGGTAGTC CTGAGACATG AGTCCGATTA CTACAGGCAC 300

GTGTTCTCCTCC AGGTGGAGGC TCAGGTCCCC GGGTGAGCTG GGGCTGCAGC GGGACTCAGG 360

GCGCGGCTCT GGCTGCAGGT CTCGCAGCTC CCTGGGCTGT AGCTCCCGCA GATCCTTGCG 420

CACACCGTTG ACTGGT 436

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

08962560-103197

- 151 -

- (A) LENGTH: 2180 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TTAATAGTAC CTACATAGTA GAAAATTATA ACTCCACTTT AAAACAATGT TTTCTTTCTA	60
TTCAAATCAA TTTAAACTT TTTATAACA TTAATGTTGC AAGAGAATCC AGTCCATTTA	120
TGAAAATTAG TTGACAATCA AGTTCACCCA AGAAAATGTT GACTAAGCTA AAGAAATCAC	180
AGATAAAACA TTTTACCAA AGGATAGGTA ACACACAAAA AAATGCTATC ACAGGAAGCT	240
ATGATCATCT AATATTTCTT TAATAATAAT TCTAGTTCCA TAGGTTTCA TGTTATGCCA	300
ATTTGTACCC GAGTTTAATT ACAGAAAAGG CAACAATTTT TAAATTGGTG GTATACATTT	360
CTTTACAATT TTTTAATGTA AGGCCATTTA TTAATAGTA CAACTAGAA GATGAAAACG	420
AAGGCAACAG AAAAATTCAA CTTTTCACAA CCAAAAGAAT TAGCACAACC TTAGAAATAA	480
TTTAGAAAAA AGTGTGTGTA AAAGATATGT TGCAGATCTC CGTTCCATTA CCCAAGATTA	540
TGTC AATTCA CGATTCTAAA TAAATCTTTT TAAAGTAAGA GATTAAAAAC TCATCTTCAG	600
TGTATATGTA AATTCCGTGG TTTATCACA CAGGTATGTT TATTCAACAC TGCTTTGGAA	660
ATGGACCATT TAAAAGGACA TGGCAATTTT CATTCTGTGA AGTTTCATTC AACCTTTACT	720
TAGGGGTTGA TTACCACATG AAATGTGCTT TTAATGCATA AAAATCACAG TGGATTAGCC	780
AGCAAAAGGG ACTGGGCGGG GGGGCGATTG AGGAGAATTT GATAATTCAC ATTGTGATTA	840
TTCTGCACAT TGATGAAACA TAATTCACAC CTCTAAAACC TCAAGACTTC CCTTTTTTAA	900
AGAACCAAAA TAAACCCAAG ACACCTTGCT GACACTTCCC CACCCCTAAA CAACTGATG	960
ACTCTTTTAC ACATAAACT GAAATAGTTA TGGCAGCAAA AGATTTTGAT GGCAATGAAA	1020

08962550-103197

- 152 -

GTTTGTAAC TGTATTCAA TCTCTTGTTT TTATCCCAA AGTGCAAGAT GCAGGGTTCT	1080
CAATCTTTCA GTAGTGCTTC TCCTGTAAAT AATCCTTCAT TTTGTTTGGC AAAGGCAGTT	1140
TCTGAATTAA GTCTATTCTG GTATACTGAC GTATAACAAA ACGACACAGG TACTGCAACG	1200
AGCGCACCTA TGAACCCCGG AACACTGGTT GGCAAGTTCT GACGGAAGTG CAGATTCCAG	1260
GCAGCGAGAC CTTGAATAAC AAAAAGCTCC CATTTTCAGA GTCCCTGATT GAATGCTCCA	1320
ATTAGATCAA CTATGGACGT ATGTCCTTCC ACATCGGCTG TTCATAAAG CTAAACCTAC	1380
CATTTGAGTG CTCAATTCTA GTGTGAAGTG TTTTACCATG GGAGCGAAAG TCACAGCTTA	1440
AAAGGTAACG GTCGTCAGAA CTGTCCCGAA CAAGAAAAGA ACCATCTGGC ACGTTTGCTA	1500
GCTTCCCTTC TGCCTCCCAA CGTGTGATTG GTCCCCAGTA CCATCCTTGC TTTGCAAGTT	1560
TTTTCAGCTC CTCTGTAAGG CTTGTCACAA CCATGGGACC ACTACTTGC ACTGAGTCAT	1620
AAACTCTTGC AACCCAGGA GCAGAGTTCG GATCAAAATT CAAATGACAG CGCATAACTT	1680
TCAGCCACGT GGGGCTTTCT GTCCAGTGAG TCCACTGAAA GTTCCCCTTT GGGATTTGGA	1740
TTATTCCTGC ATTGGAGTAA CCAATGGTGA AGATTGGAGG GACATCCATC GTGAACCCGC	1800
TCTCCGGGGT TCTGCAACAT GACTCCCGTG GTGCCAATCA ACAAGCCATT CACCGGACTG	1860
ATCCAEGAAG ATCTCTGGGG CGACAAC TAG GTCTGGTCT ACCTGACTCT CATCTCGGG	1920
GAAAGCGCGC CCTCCCACTT GAGGAGGAAC CGCAGAGACT TCCATGGGAG AAGAGCTGTC	1980
CAGACAATAG CTCCGTGATC CTTCCAAAGG ATACATCCCC TCATCTAAAG GCACAGTATA	2040
CTGAATGTAG TCCTGAGGCA TAAGTCCAAT AACGACAGGC ACATGTTTCAT CCAGGTGAAG	2100
ATGCAGGTCT CCATTATGAG AAGCCGAGCT CTTCACTGAA TTGGCTTGCT CCTGGCACGT	2160
GGTCTCAGAC TGGAGGTCGT	2180

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

00962560.103497

- 153 -

- (A) LENGTH: 2649 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GGCACGAGGC TGTGTCCAGC ACACAGAGAG GGCCCGGCCA TCTGCTTTGG TTCAGAGCCC	60
TGTGTCTGTC TGTCACCTAG ACTCTTCCTC CCGGCTCGCA GCTCACCTC CATCCTCCTT	120
ACTGGCTCCA GCATGACTCG CTTCTCTTAT GCAGAGTACT TTGCTCTGTT TCACTCTGGC	180
TCTGCACCTT CCAGGTCCCC TTCGTCTCCC GAGAACCAC CGGCCCCGCG ACCCCTGGGT	240
CTGTTCCAAG GGGTCATGCA GAAGTATAGC AGCAACCTGT TCAAGACCTC CCAGATGGCG	300
GCTATGGACC CCGTGCTGAA GGCCATCAAG GAAGGGGATG AAGAGGCCTT GAAGATCATG	360
ATCCAGGATG GGAAGAATCT TGCAGAGCCC AACAAGGAGG GCTGGCTGCC GCTCCACGAG	420
GCTGCCTACT ATGGCCAGCT GGGCTGCCTG AAAGTCCTGC AGCAAGCCTA CCCAGGGACC	480
ATTGACCAAC GCACACTGCA GGAAGAGACA GCATTATACC TGGCCACATG CAGAGAACAC	540
CTGGATTGCC TCCTGTCGCT GCTCCAGGCG GGGGCAGAGC CTGACATCTC TAACAAATCC	600
AGGGAGACTC CACTTTACAA AGCCTGTGAG CGCAAGAACG CGGAGGCGGT GAGGATATTG	660
GTGCATACA ACGCAGACGC CAACCACCGC TGTAACAGGG GCTGGACCGC ACTGCACGAG	720
TCTGTCTCCC GCAATGACCT GGAGGTCATG GAGATCCTAG TGAGTGGCGG GGCCAAGGTG	780
GAGGCCAAGA ATGTCTACAG CATCACCCCT TTGTTTGTGG CTGCCCAGAG TGGGCAGCTG	840
GAGGCCCTGA GGTTCCTGGC CAAGCATGGT GCAGACATCA ACACGCAGGC CAGTGACAGT	900
GCATCAGCCC TCTACGAGGC CAGCAAGAAT GAGCATGAAG ACGTGGTAGA GTTCTTCTC	960
TCTCAGGGCG CCGATGCTAA CAAAGCCAAC AAGGACGGCC TGCTCCCCCT GCATGTTGCC	1020

09962560-103197

- 154 -

TCCAAGAAGG GCAACTATAG AATAGTGCAG ATGCTGCTGC CTGTGACCAG CCGCACGCGC 1080  
 GTGCGCCGTA GCGGCATCAG CCCGCTGCAC CTAGCGGCCG AGCGCAACCA CGACGCGGTG 1140  
 CTGGAGGCGC TGCTGGCCGC GCGCTTCGAC GTGAACGCAC CTCTGGCTCC CGAGCGCGCC 1200  
 CGCCTCTACG AGGACCGCCG CAGTTCTGCG CTCTACTTCG CTGTGGTCAA CAACAATGTG 1260  
 TACGCCACCG AGCTGTTGCT GCTGGCGGSC GCGGACCCCA ACCGCGATGT CATCAGCCCT 1320  
 CTGCTCGTGG CCATCCGCCA CGGCTGCCTG CGCACCATGC AGCTGCTGTT GGACCATGGC 1380  
 GCCAACATCG ACGCCTACAT CGCCACTCAC CCCACCGCCT TTCCAGCCAC CATCATGTTT 1440  
 GCCATGAAGT GCCTGTCGTT ACTCAAGTTC CTTATGGACC TCGGCTGCGA TGGCGAGCCC 1500  
 TGCTTCTCCT GCCTGTACGG CAACGGGCCG CACCACCCGC CCCGCGACCT GGCCGCTTCC 1560  
 ACGACGCACC CGTGGACGAC AAGGCACCTA GCGTGGTGCA GTTCTGTGAG TTCCTGTCGG 1620  
 CCCCAGGAAGT GAGCCGCTGG GCGGGACCCA TCATCGATGT CCTCCTGGAC TATGTGGCA 1680  
 ACGTGCAGCT GTGCTCCCGG CTGAAGGAGC ACATCGACAG CTTTGAGGAC TGGGCTGTCA 1740  
 TCAAGGAGAA GGCAGAACCT CCGAGACCTC TGGCTCACCT CTGCCGCTG CGGGTTCCGA 1800  
 AGGCCATAGG AAAATACCGG ATAAACTCC TGGACACACT GCCGCTTCCC GGCAGGCTAA 1860  
 TCAGATACTT GAAATATGAG AATACACAGT AACCAGCCTG GAGAGGAGAT GTGGCCTTCA 1920  
 GACTGTTTCC GGGACGCCCC AGGTGGCCTG CATCCAGGAC CCCCTGGGGT CAGAACAGGT 1980  
 GTGACCTTGC TGGTCTTTG CTGGAGCTTC ACCCAAAGTG AGAACCTGAT GTGGGGAGTG 2040  
 GACGTGGAAC CTCTGCTTTC AACTGTCTAG CGGATCGCAG ACCCGCTCTG CTTCTGGCCA 2100  
 TAGCCAGAGA CTTTCAACCT GGGGCCAGGG GAGAGCTGCT CTGGGCAAGG TGGCCCAGGC 2160  
 AGGAATCCTG GCCTTAAGCT GGAGAACTTG TAGGAATCCC TCACTGGACC CTCAGCTTTC 2220  
 AGGCTGCGAG GGAGACGCCC AGCCCAAGTA TTTATTTTCC GTGACACAAT AACGTTGTAT 2280  
 CAGAAAAAAA AAAAAACATG GGCGCAGCTT ATTCCTTAGT AGGGTATTTA CTTGCATGCG 2340  
 CGCTTAAAGC TACTGGAAAC ATGCGTTCCA CTATGCTTGA GAATCCCTT GCACTGTAA 2400

08962560 - 103197



- 155 -

ACGAGAGCCG ACGTGCTTCA AGGTTGGATT TTTGGTTGCC CCTTTGGCGT TCCGCGGGTT	2460
TGTCCGACGT AATTGACCCC GTGTTTGTG ACTTTCGAGT GTTCCGACTA TTGGGGGGCT	2520
TTTGGTTGTC CCCAAAATTG TGGGTGGTGT GCGGACGCCA CGAGAAGTGG TTCATGGGCG	2580
ATAATCATTA CTGGAGAATG TAGAGCGGCG GTTTTACGAA TAAATATTTT TTAAGCCGCC	2640
TTCCCAAAA	2649

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 495 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CCTCCTGAGA GTTCGCCGGC CCGGGCCCAA TGGGTTGTTC CAAGGGGTCA TGCAGAAATA	60
CAGCAGCAGC TTGTTCAAGA CCTCCAGCT GCGCCTGCG GACCCCTTGA TAAAGGCCAT	120
CAAGGATGCG ATGAAGAGGC CTTGAAGACC ATGATCAAGG AAGGAAGAA TCTCGCAGAG	180
CCCAACAAGG AGGGCTGGCT GCCGCTGCAC GAGGCCGCAT ACTATGGCCA GGTGGGCTGC	240
CTGAAAGTCC TGCAGCGAGC GTACCCAGGG ACCATCGACC AGCGCACCTT GCAGGAGGAA	300
ACAGCCGTTT ACTTGGCAAC GTGCAGGGGC CACCTGGACT GTCTCCTGTC ACTGCTCCAA	360
GCAGGGGCAG AGCGGGACAT CTCCAACAAA TCCCGAGAGA ACCGCTCTAC AAAGCCTGTG	420
AGCGCAAGAA CGCGGAAGCC GTGAAGATTC TTGGTGCAGC ACAACGCAGA CACCAACAAC	480
GCTGCAACCG GGCTG	495

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

00962560-103197

- 156 -

- (A) LENGTH: 709 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTGCAGCTCT GCTCGCGGCT GAAGGAACAC ATCGACAGCT TTGAGGACTG GGCCGTCATC	60
AAGGAGAAGG CAGAACCTCC AAGACCTCTG GCTCACCTTT GCCGACTGCG GGTTCGAAAG	120
GCCATTGGGA AATACCGTAT AAAACTCCTA GACACCTTGC CGCTCCCAGG CAGGCTGATT	180
AGATACCTGA AATACGAGAA CACCCAGTAA CTGGGGCCAC GGGGAGAGAG GAGTAGCCCC	240
TCAGACTCTT CTTACTAAGT CTCAGGACGT CGGTGTTCCC AACTCCAAGG GGACCTGGTG	300
ACAGACGAGG CTGCAGGCTG CCTCCCTCTC AGCCTGGACA GCTACCAGGA TCTCACTGGG	360
TCTCAGGGCC CAGAGCTTTG GCCAGAGCAG AGAACAGAAT GTGTCAAGGA GAAGAATCAT	420
TTGTTTACAA ACTGATGAGC AGATCCCAGA CCTTCTCTAC CTTCAGGAAT GGCAGAAACC	480
TCTATTCTTG GGGCCAGGGC AGAGCTTGAG GTGTCTGGG GAAGGTGGTG CTCAGAGCCT	540
-----	
TCCCTGTGCC CCTCCACTG TTCTGGAAAA CTCACCACTT GACTTCAGAG CTTTCTCTCC	600
AAAGACTAAG ATGAAGACGT GGCCCAAGGT AGGGGGTAGG GGGAGCCTGG GTCTTGAGG	660
GCTTTGTAA GTATTAATAT AATAAATGTT ACACATGTGA AAAAAAAAAA	709

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 848 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

03952560-103197

- 157 -

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TTG GAG AAG TGT GGT TGG TAT TGG GGG CCA ATG AAT TGG GAA GAT GCA	48
Leu Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala	
1 5 10 15	
GAG ATG AAG CTG AAA GGG AAA CCA GAT GGT TCT TTC CTG GTA CGA GAC	96
Glu Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp	
20 25 30	
AGT TCT GAT CCT CGT TAC ATC CTG AGC CTC AGT TTC CGA TCA CAG GGT	144
Ser Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly	
35 40 45	
ATC ACC CAC CAC ACT AGA ATG GAG CAC TAC AGA GGA ACC TTC AGC CTG	192
Ile Thr His His Thr Arg Met Glu His Tyr Arg Gly Thr Phe Ser Leu	
50 55 60	
TGG TGT CAT CCC AAG TTT GAG GAC CGC TGT CAA TCT GTT GTA GAG TTT	240
Trp Cys His Pro Lys Phe Glu Asp Arg Cys Gln Ser Val Val Glu Phe	
65 70 75 80	
ATT AAG AGA GCC ATT ATG CAC TCC AAG AAT GGA AAG TTT CTC TAT TTC	288
Ile Lys Arg Ala Ile Met His Ser Lys Asn Gly Lys Phe Leu Tyr Phe	
85 90 95	
TTA AGA TCC AGG GTT CCA GGA CTG CCA CCA ACT CCT GTC CAG CTG CTC	336
Leu Arg Ser Arg Val Pro Gly Leu Pro Pro Thr Pro Val Gln Leu Leu	
100 105 110	
TAT CCA GTG TCC CGA TTC AGC AAT GTC AAA TCC CTC CAG CAC CTT TGC	384
Tyr Pro Val Ser Arg Phe Ser Asn Val Lys Ser Leu Gln His Leu Cys	
115 120 125	
AGA TTC CGG ATA CGA CAG CTC GTC AGG ATA GAT CAC ATC CCA GAT CTC	432
Arg Phe Arg Ile Arg Gln Leu Val Arg Ile Asp His Ile Pro Asp Leu	
130 135 140	

00962560-103197

CCA CTG CCT AAA CCT CTG ATC TCT TAT ATC CGA AAG TTC TAC TAC TAT 480  
Pro Leu Pro Lys Pro Leu Ile Ser Tyr Ile Arg Lys Phe Tyr Tyr Tyr  
145 150 155 160

GAT CCT CAG GAA GAG GTA TAC CTG TCT CTA AAG GAA GCG CAG CGT CAG 528  
Asp Pro Gln Glu Glu Val Tyr Leu Ser Leu Lys Glu Ala Gln Arg Gln  
165 170 175

TTT CCA AAC AGA AGC AAG AGG TGG AAC CCT CCA CGT AGC GAG GGG CTC 576  
Phe Pro Asn Arg Ser Lys Arg Trp Asn Pro Pro Arg Ser Glu Gly Leu  
180 185 190

CCT GCT GGT CAC CAC CAA GGG CAT TTG GTT GCC AAG CTC CAG CTT TGAAGAACCA  
631  
Pro Ala Gly His His Gln Gly His Leu Val Ala Lys Leu Gln Leu  
195 200 205

AATTAAGCTA CCATGAAAAG AAGAGGAAAA GTGAGGGAAC AGGAAGGTTG GGATTCTCTG 691

TGCAGAGACT TTGGTTCCCC ACGCAAGCCC TGGGGCTTGG AAGAAGCACA TGACCGTACT 751

CTGCGTGGGG CTCCACCTCA CACCCACCCC TGGGCATCTT AGGACTGGAG GGGCTCCTTG 811

GAAAACCTGGA AGAAGTCTCA ACACTGTTTC TTTTCA 848

(2) INFORMATION FOR SEQ ID NO:36:

—(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Leu Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala  
1 5 10 15

Glu Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp  
20 25 30

Ser Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly  
35 40 45

- 159 -

Ile Thr His His Thr Arg Met Glu His Tyr Arg Gly Thr Phe Ser Leu			
50	55	60	
Trp Cys His Pro Lys Phe Glu Asp Arg Cys Gln Ser Val Val Glu Phe			
65	70	75	80
Ile Lys Arg Ala Ile Met His Ser Lys Asn Gly Lys Phe Leu Tyr Phe			
85	90	95	
Leu Arg Ser Arg Val Pro Gly Leu Pro Pro Thr Pro Val Gln Leu Leu			
100	105	110	
Tyr Pro Val Ser Arg Phe Ser Asn Val Lys Ser Leu Gln His Leu Cys			
115	120	125	
Arg Phe Arg Ile Arg Gln Leu Val Arg Ile Asp His Ile Pro Asp Leu			
130	135	140	
Pro Leu Pro Lys Pro Leu Ile Ser Tyr Ile Arg Lys Phe Tyr Tyr Tyr			
145	150	155	160
Asp Pro Gln Glu Glu Val Tyr Leu Ser Leu Lys Glu Ala Gln Arg Gln			
165	170	175	
Phe Pro Asn Arg Ser Lys Arg Trp Asn Pro Pro Arg Ser Glu Gly Leu			
180	185	190	
Pro Ala Gly His His Gln Gly His Leu Val Ala Lys Leu Gln Leu			
195	200	205	

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

00962560-103197

- 160 -

GTTCCAAGCC TAACCCATCT TTGTCGTTTG GAAATTCGGG CCAGTCTAAA AGCAGAGCAC 60

CTTCACTCTG ACATTTTCAT CCATCAGTTG CCACTTCCCA GAAGTCTGCA GAACTATTTG 120

CTCTATGAAG AGGTTTTAAG AATGAATGAG ATTCTAGAAC CAGCAGCTAA TCAGGATGGA 180

GAAACCAGCA AGGCCACCTG ACACAGGTCC TTTAATTCTG TTTAGTCACA AAAGACGGCT 240

TGTGTGACTG TTTGGATTTG GTGATCAAAT GTCCATGTTT ACAGTTGCTT TTCCAGTTT 300

GTGTCTTTCC CAATATTGTG AACCTTATCC ATCTTGCCTT ACTCAGTTT ATTTCTAGTG 360

CACTTGTTG TGTATTATTT GTTTACCTGA CCATTTTCTA CTTTATTCTG CTAATAAACT 420

GTAATTCTGA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA 464

## (2) INFORMATION FOR SEQ ID NO:38:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 747 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGGGATCGAA AGCGGGGGCT TCTGGGACGC AGCTCTGGAG ACGCGGCCTC GGACCAGCCA 60

TTTCGGTGTA GAAGTGGCAG CACGGCAGAC TGGTCAAACA AATGGATTTT ACAGAGGCTT 120

ACGCGGACAC GTGCTCTACA GTTGACTTG CTGCCAGGGA AGGCAATGTT AAAGTCTTAA 180

GGAAACTGCT CAAAAAGGGC CGAAGTGTCT ATGTTGCTGA TAACAGGGGA TGGATGCCAA 240

TTCATGAAGC AGCTTATCAC AACTCTGTAG AATGTTTGCA AATGTTAATT AATGCAGATT 300

CATCTGAAAA CTACATTAAG ATGAAGACCT TTGAAGGTTT CTGTGCTTTG CATCTCGCTG 360

CAAGTCAAGG ACATTGGAAG ATCGTACAGA TTCTTTTAGA AGCTGGGGCA GATCCTAATG 420

CAACTACTTT AGAAGAAACG ACACCATTGT TTTTAGCTGT TGAAAATGGA CAGATAGATG 480

08962560-103197

- 161 -

TGTTAAGGCT GTTGCTTCAA CACGGAGCAA ATGTTAATGG ATCCCATTCT ATGTGTGGAT	540
GGAACTCCTT GCACCAGGCT TCTTTTCAGG AAAATGCTGA GATCATAAAA TTGCTTCTTA	600
GAAAAGGAGC AAACAAGGAA TGCCAGGATG ACTTTGGAAT CACACCTTTA TTTGTGGCTG	660
CTCAGTATGG CCAAGCTAGA AAGCTTTGAA GCATACTTAT TTCATCCGGG TGCAAATGTC	720
AATTGTCAAG CCTTGGACAA AGCTACC	747

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1018 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CACAAATGGG ACCATACAAA AATCTTGGAC TTGTTAATAA CCACTTACTA ACCGGGACCT	60
GTGACACTGG GCTAAACAAA GTAAGTCCCT GTTACTCAG CAGTGTTTGG GGGACATGAA	120
GGATTGCTTA GAAATATTAC TCCGGAATGG TCTACAGCCC AGACGCCAG CCGTGCCTTG	180
TTTTTGATT CAGTTCTCCT GTGTGCATGG CTTTCCAAA GGAGGTGGAG CTGTAGTTCT	240
TTGGAATTGT GAACATTCTT TTGAAATATG GAGCCCAGAT AAATGAACTT CATTGGCAT	300
ACTGCCTGAA GTACGAGAAG TTTTCGATAT TTCGCTACTT TTTGAGGAAA GGTGCTCAT	360
TGGGACCATG GAACCATATA TATGAATTG TAAATCATGC AATTAAAGCA CAAGCAAAAT	420
ATAAGGAGTG GTTGCCACAT CTTCTGGTTG CTGGATTGTA CCCACTGATT CTA CTGTGCA	480
ATTCTTGGAT TGACTCAGTC AGCATTGACA CCCTTATCTT CACTTTGGAG TTTACTAATT	540
GGAAGACACT TGCACCAGCT GTTGAAAGGA TGCTCTCTGC TCGTGCCTCA AACGCTTGA	600
TTCTACAGCA ACATATTGCC CACTGTTCCA TCCCTGACCC ATCTTTGTCT TTTGGAAATT	660

08962560 103197

- 162 -

CGGTCCAGTC TAAATCAGA ACGTCTACGG TCTGACAGTT ATATTAGTCA GCTGCCACTT	720
CCCAGAAGCC TACATAATTA TTTGCTCTAT GAAGACGTTT TGAGGATGTA TGAAGTTCCA	780
GAAGTGGCAG CTATTCAAGA TGGATAAATC AGTGAAACTA CTTAACACAG CTAATTTTTT	840
TCTCTGAAAA ATCATCGAGA CAAAAGAGCC ACAGAGTACA AGTTTTTATG ATTTTATAGT	900
CAAAAGATGA TTATTGATTG TCAGATAGGT TAGGTTTTGG GGGGCCAGTA GTTCAGTGAG	960
AATGTTTATG TTTACAAC TA GCCTTCCCAG TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1018

## (2) INFORMATION FOR SEQ ID NO:40:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1897 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CGGGGGGCTG GGACCTGGGG CGTAACCGTC TCTACCACGA CGCAAGAAC CAGCCAAGTA	60
AAACATACCC AGCCTTTCTG GAGCCGGACG AGACATTCAT TGTCCCTGAC TCCTTTTTCG	120
TGGCCCTGGA CATGRATGAT GGGACCTTAA GTTTCATCGT GGATGGACAG TACATGGGAG	180
TGGCTTTCCG GGGACTCAAG GGTAAAAAGC TGTATCCTGT AGTGAGTGCC GTCTGGGGCC	240
ACTGTGAGAT CCGCATGCGC TACTTGAACG GACTTGATCC TGAGCCCCTG CCACTCATGG	300
ACCTGTGCCG GCGTTCGGTG CGCCTAGCGC TGGGAAAAGA GCGCCTGGGT GCCATCCCCG	360
CTCTGCCGCT ACCTGCCTCC CTCAAAGCCT ACCTCCTCTA CCAGTGATCC ACATCCCAGG	420
ACCGCCATAC GACAGCCATC TGGTGCCAAR TCACTGAGCC CGTTGGGGTC CGCCGACCCC	480
TGCGCCTGGG ATGGAAGCCC ACCTCAGCCA TGGGCAGACG TGCCCCCTCA TCCTACCGGC	540

09962560-103197



- 163 -

TGCCTCTGCT GGGGGAACCT ATGCCAACGG ACTTCTCCCT TCCCAACACT GGCTGAAGCA 600  
 GCAGCACCCA GGGCCTTCCC TGAACCAGAT GCAGAGAATA AACTATGAAA ACCTCTCTCA 660  
 GGCGCCTTCT GCTCTCAGGT GGAGTGGGCT GGGGGGCACT CTCTGCAGAG AGAGGCTACA 720  
 CCCACCTGGG GGGTCCTGGG AGGTAAGACT AGTAGGAGGT GCCAGGGCTG ARTCCAAAAG 780  
 CAGGAATGGC CAGGAMCAGG CCATACAGAT GAAGCTCAGG ATGTCACATA CCATGGACAM 840  
 TGAGACAGAA CCCCAGGTTG GAMTTCCCTT GGGCCAACGA GTGCCAGCTT TAATGTCAGC 900  
 TCGMGGTGCT CTGTGGCCTG TATTTATTCT TTAAACAGTA GCAAAGGCCA TTTATTTATT 960  
 CCACTTAGAA AGGAAACCTT GGTGGGTGGY TTCCCTCGAT GTGCTTTCCC CCACCTCCCT 1020  
 GGAATGTGTG TGCCACACCT GTCCTTGTC CAGGCCAGGA CTGTGGCACA TGAGCTGGTG 1080  
 TGCACAGATA CACGTATGTC GTCGTGCATG ACCCCTGACT AGTTCCTAAG TAGCCCTGCA 1140  
 CCAAGCACCA GAGCAGACCC CAAGAGAGGC CCGTGCAAGT CCCCATGTCC CCAGGTCCCT 1200  
 GCTTCTGTTG CTTTGGGACT CATAACCGG CACACGTGTT TCAGCCTCTT GACTTCCATG 1260  
 AGCTTCGAAT TTGCCCCCG ATTCTTCTGA TATTTCCCAT TGGCATCCTC CAAAGCTCTG 1320  
 GGCTGGAGG GCATTAGGAC ACATGGAATG AGTGGGGTCT CCAGCCCCTG GGAAAGCCAC 1380  
 TGGCAAGGCA GGATTAGAAA GACCAAGAGC AGGGTGGGGC GCCATGAAGC CTGTATGCCT 1440  
 CTCAGGCTCA AGACCCCGCC ACACACCCAC TCAAGCCTCA GAAGTGGTGT GTAGGGCAGC 1500  
 CCCAGGAGAG GAATGCCTGT CCTAGCAGCA CGTACATGGA GCACCCACAC TGTGCTCCAG 1560  
 CCTCTGGCT GTTTCTCTTG CTCTAGAATC AACTCCCTAC ATTGGGAATG TAGCCATTTG 1620  
 GTAGAGGACT TGCCTAGCCT GCAGGAAGCT CACGTTCCAT CCCCTGCACC AAGGAGAATC 1680  
 AAAGCTCAGG AGGCTGAGGC AGCAGGATTG CTGTCAGTGG TGTACAGAGG TCATGGCCAT 1740  
 CCTGGGCTAT ATTAAACCTT GTCCTTTAAG AAAAAGAAAA GAAATCAACT TCCATTGAAT 1800  
 CTGAGTTCTG CTCATTTCTG CACAGGTACA ATAGATGACT TKATTTGTTG AAAAATGKTT 1860  
 AATATATTTA CMTATATATA TATTTGTAAG AAGCATT 1897

00962560-103197

(i) SEQUENCE CHARACTERISTICS:

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ala Tyr Leu Leu Tyr Gln  
130

(2) INFORMATION FOR SEQ ID NO:42:

**090607Z JUL 80**

- 165 -

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAGGGTAAAA AACTGTATCC TGTAGTGAGT GCCGTCTGGG GCCACTGTAG ATCCGAATGC 60  
GCTACTTGAA CGGACTCGAT CCCGAGACTG CCGCTCATGG ATTTGTGCCG TCGCTCGGTG 120  
CGCCTGGCCC TGGGGAGGGA GCGCCTGGGG GAGAACCACA CCTGCCGCTG CCGGCTTCCC 180  
TCAAGGCCTA CCTCCTCTAC CAGTGACGTT CGCCATCATA CCGCCAGCGC GACAGCCACC 240  
TGGTGCCAAC TCACTGAGCC GCCTG 265

## (2) INFORMATION FOR SEQ ID NO:43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2438 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AAGTGCGGC GGTCCCTGGA GAGCAGGCGG AGGCAGCGGC AAGTCTGACT CTGGGCTGAC 60  
CGTGGAGCCG GGGCGGGGGC TGACAGCCAG GCCTCCGCCT GGCGGGAGCC GCACGAGGAG 120  
CGGGACTGGC CGGGCCTCTC TTCCGCGCTT GAGCGAGCGC CGGGTGATGG CGGTGGTGAT 180  
GGCGGCAGGC GCTCGGACAG CTCCGCTTGA GCTGAGCTCG GAGAGATCCG TCCAGAAAGT 240

26 FEB - 09 52 56 80

- 166 -

GCCCAGAAGA AACTTCCTCT TAGAAAAGCT GAAAAACACA RTATTTATAA CACTGGAAAT	300
TGTAAAGAAT TTGTTTAAAA TGGCTGAAAA CAATAGTAAA AATGTAGATG TACGGCCTAA	360
AACAAGTCGG AGTCGAAGTG CTGACAGGAA GGATGGTTAT GTGTGGAGTG GAAAGAAGTT	420
GTCTTGGTCC AAAAAGAGTG AGAGTTGTTC TGAATCTGAA GCCATAGGTA CTGTTGAGAA	480
TGTTGAAATT CCTCTAAGAA GCCAAGAAAG GCAGCTTAGC TGTCGTCCA TTGAGTTGGA	540
CTTAGATCAT TCCTGTGGGC ATAGATTTTT AGGCCGATCC CTAAACAGA AACTGCAAGA	600
TGCGGTGGGG CAGTGTTTTT CAATAAAGAA TTGTAGTGGC CGACACTCTC CAGGGCTTCC	660
ATCTAAAAGA AAGATTCATA TCAGTGAACCT CATGTTAGAT AAGTGCCCTT TCCCACCTCG	720
CTCAGATTTA GCCTTTAGGT GGCATTTTAT TAAACGACAC ACTGTTCTTA TGAGTCCCAA	780
CTCAGATGAA TGGGTGAGTG CAGACCTGTC TGAGAGGAAA CTGAGAGATG CTCAGCTGAA	840
ACGAAGAAAC ACAGAAGATG ACATACCCTG TTTCTCACAT ACCAATGGCC AGCCTTGTGT	900
CATAACTGCC AACAGTGCTT CGTGTACAGG TGGTCACATA ACTGGTTCTA TGATGAACTT	960
GGTCACAAAC AACAGCATAG AAGACAGTGA CATGGATTCA GAGGATGAAA TTATAACGCT	1020
GTGCACAAGC TCCAGAAAAA GGAATAAGCC CAGGTGGGAA ATGGAAGAGG AGATCCTGCA	1080
GTTGGAGGCA CCTCTAAGT TCCACACCCA GATCGACTAC GTCCACTGCC TTGTTCCAGA	1140
CCTCCTTCAG ATCAGTRACA ATCCGTGCTA CTGGGGTGTC ATGGACAAAT ATGCAGCCGA	1200
AGCTCTGCTG GAAGGAAAGC CAGAGGGCAC CTTTTACTT CGAGATTCAG CGCAGGAAGA	1260
TTATTTATTC TCTGTTAGTT TTAGACGCTA CAGTCGTTCT CTTCATGCTA GAATTGAGCA	1320
GTGGAATCAT AACTTTAGCT TTGATGCCCA TGATCCTTGT GTCTTCCATT CTCCTGATAT	1380
TACTGGGCTC CTGGAACACT ATAAGGACCC CAGTGCCTGT ATGTTCTTTG AGCCGCTCTT	1440
GTCCACTCCC TTAATCCGGA CGTTCCCTT TCCCTGTCAG CATATTTGCA GAACGGTTAT	1500
TTGTAATTGT ACGACTTACG ATGGCATCGA TGCCCTTCCC ATTCCTTCGC CTATGAAATT	1560
GATCTGAAG GAATACCATT ATAAATCAAA AGTTAGGTTA CTCAGGATTG ATGTGCCAGA	1620

00962560-103197

- 167 -

GCAGCAGTGA TGCGGAGAGG TTAGAATGTC GACCTGCATA CATATTTTCA TTAAATATTT 1680  
TATTTTCTT ATGCTCTTT GAATTTTGT ACAAAGGCAG TTGAATCAAA TAAACTGTG 1740  
CCCTAAGTTT TAATTCCAGA TCAATTTATT TTTTATGA TACACTTGT ATATATTTT 1800  
AAGCAGGTGT TTGGTTTGT TTTTACCATA TAAATTTACA TATGGTCCAG GCATATTAC 1860  
AATTCAAGG CATTGCATAT ACATTTGAAT ATTCTGTATT TTTTAAATA TCTTTGTTC 1920  
TTTCCTATGT GTGAATATT TTGCTAATCT ATGCTATCAG TATTCTTGT TGACCGAATA 1980  
GTTACCTATT CTCTTTTCAT CTGAAGATT TTCAGTAAAG AGTGTGTAA TCAATCCATT 2040  
ATAATGTAAT TGACTTTGT AATTGCCAA TAGGAGTGT AAACAACAAA ATGATTTAA 2100  
ATGAACTTA ATGTATTTT ATTTTAAATA TTAATAAAC CAAGTTGTG TGTAGTTAT 2160  
TCTAGCCAAT AAGAAAGAG AATGTAGCAT CCTAGAGGTG TATTGTCT GCAGTTGGC 2220  
AGGACCGTCA GTTAGTCCAA ATAAACATCC CCTCAGCGTG GAGGCGAATG GAACCTGTGC 2280  
TCCTTTCTTA CGGGAAGCTT TGCAAAGCAA AATAGCAGG TTACACCTT GGAGTTGTTA 2340  
AGGCAACTAG AGTTTCTCT ATTAATTAT AGACTGTTGT TGCACCTACT TAGCTCTTT 2400  
TTGGGAAGTC TAGTCCAG GGGAAATAC CTCGTGCC 2438

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 542 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ser Gly Gly Gly Pro Trp Arg Ala Gly Gly Gly Ser Gly Lys Ser Asp  
1 5 10 15

09962560-103197

- 168 -

Ser Gly Leu Thr Val Glu Pro Gly Arg Gly Leu Thr Ala Arg Pro Pro  
20 25 30

Pro Gly Gly Ser Arg Thr Arg Ser Gly Ser Gly Arg Ala Ser Leu Pro  
35 40 45

Arg Leu Ser Glu Arg Arg Val Met Ala Val Val Met Ala Ala Gly Ala  
50 55 60

Arg Thr Ala Pro Leu Glu Leu Ser Ser Glu Arg Ser Val Gln Lys Val  
65 70 75 80

Pro Arg Arg Asn Phe Leu Leu Glu Lys Leu Lys Asn Thr Xaa Phe Ile  
85 90 95

Thr Leu Glu Ile Val Lys Asn Leu Phe Lys Met Ala Glu Asn Asn Ser  
100 105 110

Lys Asn Val Asp Val Arg Pro Lys Thr Ser Arg Ser Arg Ser Ala Asp  
115 120 125

Arg Lys Asp Gly Tyr Val Trp Ser Gly Lys Lys Leu Ser Trp Ser Lys  
130 135 140

Lys Ser Glu Ser Cys Ser Glu Ser Glu Ala Ile Gly Thr Val Glu Asn  
145 150 155 160

Val Glu Ile Pro Leu Arg Ser Gln Glu Arg Gln Leu Ser Cys Ser Ser  
165 170 175

Ile Glu Leu Asp Leu Asp His Ser Cys Gly His Arg Phe Leu Gly Arg  
180 185 190

Ser Leu Lys Gln Lys Leu Gln Asp Ala Val Gly Gln Cys Phe Pro Ile  
195 200 205

Lys Asn Cys Ser Gly Arg His Ser Pro Gly Leu Pro Ser Lys Arg Lys  
210 215 220

Ile His Ile Ser Glu Leu Met Leu Asp Lys Cys Pro Phe Pro Pro Arg  
225 230 235 240

Ser Asp Leu Ala Phe Arg Trp His Phe Ile Lys Arg His Thr Val Pro  
245 250 255

0896560-103197

91

- 169 -

Met Ser Pro Asn Ser Asp Glu Trp Val Ser Ala Asp Leu Ser Glu Arg  
 260 265 270  
 Lys Leu Arg Asp Ala Gln Leu Lys Arg Arg Asn Thr Glu Asp Asp Ile  
 275 280 285  
 Pro Cys Phe Ser His Thr Asn Gly Gln Pro Cys Val Ile Thr Ala Asn  
 290 295 300  
 Ser Ala Ser Cys Thr Gly Gly His Ile Thr Gly Ser Met Met Asn Leu  
 305 310 315 320  
 Val Thr Asn Asn Ser Ile Glu Asp Ser Asp Met Asp Ser Glu Asp Glu  
 325 330 335  
 Ile Ile Thr Leu Cys Thr Ser Ser Arg Lys Arg Asn Lys Pro Arg Trp  
 340 345 350  
 Glu Met Glu Glu Glu Ile Leu Gln Leu Glu Ala Pro Pro Lys Phe His  
 355 360 365  
 Thr Gln Ile Asp Tyr Val His Cys Leu Val Pro Asp Leu Leu Gln Ile  
 370 375 380  
 Ser Asn Asn Pro Cys Tyr Trp Gly Val Met Asp Lys Tyr Ala Ala Glu  
 385 390 395 400  
 Ala Leu Leu Glu Gly Lys Pro Glu Gly Thr Phe Leu Leu Arg Asp Ser  
 405 410 415  
 Ala Gln Glu Asp Tyr Leu Phe Ser Val Ser Phe Arg Arg Tyr Ser Arg  
 420 425 430  
 Ser Leu His Ala Arg Ile Glu Gln Trp Asn His Asn Phe Ser Phe Asp  
 435 440 445  
 Ala His Asp Pro Cys Val Phe His Ser Pro Asp Ile Thr Gly Leu Leu  
 450 455 460  
 Glu His Tyr Lys Asp Pro Ser Ala Cys Met Phe Phe Glu Pro Leu Leu  
 465 470 475 480  
 Ser Thr Pro Leu Ile Arg Thr Phe Pro Phe Ser Leu Gln His Ile Cys  
 485 490 495

00962550-103197

- 170 -

Arg Thr Val Ile Cys Asn Cys Thr Thr Tyr Asp Gly Ile Asp Ala Leu  
500 505 510

Pro Ile Pro Ser Pro Met Lys Leu Tyr Leu Lys Glu Tyr His Tyr Lys  
515 520 525

Ser Lys Val Arg Leu Leu Arg Ile Asp Val Pro Glu Gln Gln  
530 535 540

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4999 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CCCTCTGGGC AAGCCGCCCC CCCCCACCC ATCTACCACA CACACACACA CACACACACA 60

CACACATTCA GACCTTGGGG CAAAAACAAA GCAAAATAAC AACAACAAAA AACTGCCTG 120

TGGAAGTCC TTA CTTCAGG AAGGTTGGCA GATGAGGAGC AAGGGAACAT TTTATCAGGA 180

CTGCCACAAA GGAGTCTTTT TTTTAAATGG TTTTCAAGA CAGGGTTTCT CTGTATAGCC 240

CTGGCTGTCC TGGAGCTCAC TTTGTAGACC AGGCTGGCCT CGAACTCAGA AATTCGCCTG 300

CCTCTGCCTC CTGAGTGCTG GGATTAAAGG CGTGCAGCAC CATGTCCAAC TGGCATTTC 360

TCAATTAAGG TTCGTTCTT TCAGATAACT CTAGGTTCTG GGTCAAGCTG ACACAAGGCT 420

ACACAGCACA GTTTGTATGC CACATTCAGT TCAGAAGACA CCCAACCTCC CTGGAAGTGG 480

AACTTATGCA CATTGTGAG CTCCACTTG GGAGTGGGAA CCTGAACTGG GTCCTCTGCA 540

AGAGCAGCCG TGCTCTTAAC TGCTGAGCCA TTTAGCAGC CTCACATCAG AATTAAGTTA 600

08962560-103197

61



- 171 -

GAAATTAGCCG GGTATGAATC ATACCCTTAG AATCCTAGCA TCTGAAAGCA GAGCTAAGAG 660  
AAACAGGGAT TCAAGACCAG CTCTTGGCTA CAGAGCCCGT CCTGTCCTAG GATGGGCTAC 720  
AAGAGACTAT TTCAAAGCCA TCCAAACAAC AATAACTACA ACAACAACAA GGTAAAAATT 780  
AGGCTGGGCA CAGGGTACAC ACCTTTAATG CCAACACTCA GGAGGCAGAG GCAGGCTGAT 840  
CAGTGTGAGT TTGAGTTCAA CGTGGTCTAC ATAGGGAGTT CTAGGCCAGC AGAGGTTACA 900  
GTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCACACA CACACACACA CACACACACA 960  
CACACACACA CACACACGGT GGCATTATGG GATTTTTTTG GGATAAGGTT TCTCTGTCTA 1020  
GCCCTGGCAT AGATTCACCTC TGTAAGTAGT GCTAGCCTTG AACTCAGAGA TCCGCCTGCC 1080  
TCTGCCTCCC AAGTGCTGGG ATTATAGGTG TTGCACCACC ACTGCCCAGC CACTTTGGGA 1140  
TTTTTGAAGT GTTATCAAGA GGCTTTTCGAG GAGGTCAAAC TTCAACAGCA ACCTCTCCAT 1200  
GATAATGTAG CTAATGATCA AACGACACTC AAAACTTAAC CCTTAAAGCA CACATCCACC 1260  
AGACAGCGTG CCCACTCGTA GTTCCATTAC TCAGGAGGCT GAAGCAGGAG GATGAAGGAC 1320  
TAAGGCTTCA GCAACCTAGG GAGCCGCAGG GGACAGTAGT CTCAATCCCT ACATTCTCCT 1380  
GAACACAGGA GCAGGAGTTC AGGAAGGGTG TCAAGGCCGC TTAGTGATCT TAGGGCTCA 1440  
GGAATGACTA GCTCAGGCAG AGAGAGCAAA GGTCTCCAGT GGAGAAGTCT ACACACACAC 1500  
ACACACACAC ACACACACAC ACACACACAC AGAATCCAAG GCGATGACGT CATCAAAGGG 1560  
TTAATTCTAG TCTGGGATGG GGGGGAGGGT GGGGCACGCA GCTGTCAGGT GGCTTTGGAA 1620  
AAATAAACTG CTGAAGAGTC TGACGCCAGG GAGTCCTGGG AGGGACAAGA GGTTACCCAC 1680  
TCAAAGAGTG TGCTCCACAA AGCATGCGCG CTTGTCCACG TCTGGAGTCG TCACTTATTT 1740  
TTGCCTGGA TTCTTTCTAG CCGGTGGGTT CTCAAGGCGG TAAGTGGTGT GGCCGCCGTG 1800  
GTCTGGGAGG TGACGATAGG GTTAATCGTC CACAGAGCCC AGGGGCGGAG CGCGGGCGGG 1860  
CGTCCGCAGC CCCGCTGGAG CCGGAAGCAG TGGCTGGTCA GGGGCGCTTC TAGCCTTCCC 1920  
TATCTGTACT TCCACAGAGG TCTCTGCGAG CTAGGGGGAC AGTGAGGTGC GGGGTAGGGG 1980

00962560-103197

CCCCGCGTTA GAGCCAGCAA GGGGACGGTT CACGGTAAGG TCTGAGGGAG AGAGAGCTCC	2040
TGAGAACTT GGGGGGCGCG ACACAGATAG GGTGAAAGCA GAGTGATAGA CCTGGGATGG	2100
TTAGGGGACC AAGGGAAGAC CAGGCTGGTT GGCATACACC GGTGAACGGA TGGGAGTCCT	2160
AGGGAAAGAT GATGCGCCTA ACAGTCCTTT CTGTCTCCAC ACCACTCCAG GGGACGATCC	2220
GGAGCTCAAC TTTCAAAGC GAGACGCCCC AGCAAGCCTG TTTTGAGAAG TTCTTCAGCG	2280
GCTCTCCTCA TGGGCCAGAC GGCCCTGGCA AGGGGCAGCA GCAGCACCCC TACCTCGCAG	2340
GCTCTGTACT CGGACTTCTC TCCTCCCGAG GGCTTGAGG AGCTCCTGTC TGCTCCCCCT	2400
CCTGACCTGG TTGCCAAGC GCACCACGGC TGAACCCCA AGGATTGCTC CGAGAACATC	2460
GATGTCAAGG AAGGGGGTCT GTGCTTTGAG CGGCGCCCTG TGGCCCAGAG CACTGATGGA	2520
GTCCGGGGGA AACGGGGCTA TTCGAGAGGT CTGCACGCCT GGGAGATCAG CTGGCCCCTG	2580
GAGCAAAGGG GCACACACGC CGTGGTGGGC GTGGCCACCG CCCTCGCCCC GCTGCAGGCT	2640
GACCACTATG CGGCGCTTTT GGGCAGCAAC AGCGAGTCCT GGGGCTGGGA TATTGGGCGG	2700
GGAAAATTGT ATCATCAGAG TAAGGGCCTC GAGGCCCCC AGTATCCAGC TGGACCTCAG	2760
GGTGAGCAGC TAGTGGTGCC AGAGAGACTG CTGGTGGTTC TGGACATGGA GGAGGGGACT	2820
CTTGCTACT CTATTGGGGG CACGTACCTG GGACCAGCCT TCCGTGGACT GAAGGGGAGG	2880
ACCCCTCTATC CCTCTGTAAG TGCTGTTTGG GGCCAGTGCC AGGTCCGCAT CCGCTACATG	2940
GGCGAAAGAA GAGGTGAGAT ACGGACTAGG TGTGGGGAGA TCACTACTCT TGGCAATGGT	3000
TTGGGCTGGA AACTCATGGT TGGAGCACAG GAAGTAGGCT TCTTGCTACT TTGGCCTGTC	3060
ACTTAGATGG CCTTGGATCT AGCTTCACTC CCAATCCCTA TTGGATGTGA TGCACAAATT	3120
CAGAGCCTTT GGGTCTCCCT CAGCTGAGGT GGCGGTGGAA ATGGAGGAAG AAGGAAGGGT	3180
GCCTGAGCAG GATCTCAAGT TCAAGGATGC CTGGAGTTGC TTAATTACCT TGTCTTCCTT	3240
CTCTCTCCGC AGTGGAGGAA CCACAATCCC TTCTGCACCT GAGCCGCCTG TGTGTGCGCC	3300
ATGCTCTGGG GGACACCCGG CTGGGTCAAA TATCCACTCT GCCTTTGCCC CCTGCCATGA	3360

**THE UNIVERSITY OF CHICAGO**

- 173 -

AGCGCTATCT GCTCTACAAA TGACCCAGTA GTACAGGGTG TGCTGGCACC CTACCGTGGG 3420  
 GACAGGTGGA GAGGCACCCG CTGGCCTAGA CAACTTTAAA AAGCTGGTGA AGCTGGGGGG 3480  
 GGGGGGCTGG ACCCCTTCAC CTCCCCTTCT CACAGGAGCA AGACATATAG AAATGATATT 3540  
 AAACACCATG GCAGCCTGGG ACAAAGAGGT TTTTGAAGTA AAAAATGAGA TGTATTGTCA 3600  
 CAACCTGTTT CATTATTGTT TTTTGTTTTG TTTTACTC CCCCACCCA GGCTAGAGCC 3660  
 CCATCACTGT CTTAAGGAAT TATGACAACC CACAAAGCTC AGGCCAGGT GTTTATTTC 3720  
 CTTACATGTA GGATGGTCA CAAACACAAT ACAGGGGCTT TGGCACCCTG GGGGAGGGGA 3780  
 CTATCCCAGG CCTCTTAGGG TCTCATGTAT ACCGAATTCA GACCCGAAAG CTCTGAATTT 3840  
 CTGCATCAGA CATCCAGTAG AACTTGGGAG TGAAGCTAGA GCCAAGGCCA TCTAAGTGAC 3900  
 AGGCCAAAGT GACACGAAGC CCACTTCCTG TGCTCCAACC ATGAGTTTCC AGCCCAAACC 3960  
 AATGGAAGGT GATTTCCTT GTCAGGGCCC AAAGGGACAG TCAGTTCTAC TCCCTCCCCT 4020  
 CACTAGGAGC CACCTTGGTG ACAGTTGATT CTACCCACTG TAAGTGGTAA AGGGATTGGC 4080  
 CTGGTCCCAA CCATAATAGG GCGGTGGAAA CGGCTCAGGA GGGTACAGCG TGGATTAGGC 4140  
 CACAAGATGG GGCAGATGAT GTCATCAGAA GCATGTGACC GGTGGGAGCA GTTACTAAAC 4200  
 TTCTGGGCAA CCTAGTCCAT GCTATGCAGG CAGGTAGAGG GATGGGCAGT GCTCATTGTT 4260  
 TGGCATTGAT GATGTCCACA AATTCAGGCT TGAGAGATGC GCCACCCACA AGGAAGCCGT 4320  
 CCACGTCAGG CTGGCTTGCC AGCTCTTGC AGGTGCTCC AGTCACAGAA CCTGTACCAG 4380  
 GAACAAGAAG ACAGTTTGGT CAGGTCTATG ATCAGAACAC TTAAGCCCCA CCTCTCTGTG 4440  
 CAAGGCAGCC TCAGTCTGTC TTAGCCCATT TCCGTCTTAG CTAGAGCCAA AGCCACTCAC 4500  
 CTCCATAAAT GATCCGGGTG CTCTGAGCCA CCCCATCATT GACATTGGAT TTCAGCCATC 4560  
 CCCGGAGCTT CTCGTGTACT TCCTGTGCCT AGAAGGAGGA GGCAGAGCTA CTAAGTAAGC 4620  
 TCCTTCCTAT CTATCATCA AGGAGTAAAA ACCACTGGTT CTCACATAGA GTTGAGTTTC 4680  
 CAGAAAAGCC CCGGGACCAG AGAGTGGCAA GGCTCCAATC CCACCAGGCT TGAATGAAC 4740

0962560-103197

- 174 -

ATTTTGGCA AAGTCACTCT CCTGGTGAG TTGGGGGCC CTCTGTCTCT AAAGGGGCTT 4800  
 GGATGGGCTC CATAGCTGTG TGAGTCTGTT AAAGCCGGAC AGGCTGAGGA GCTCTGGGTA 4860  
 GTTACCTGCT GAGGGGTTCG CGTCTTGCCA GTCCCAATGG CCCACACAGG TTCATAGGCC 4920  
 AGGACCACCT TGCTCCAGTC TTTCACATTA TCTGTGGGGC AGAGAGGAGA GTGAGTAGGA 4980  
 AGGAGCTGAC CCGCCAAGC 4999

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met Gly Gln Thr Ala Leu Ala Arg Gly Ser Ser Ser Thr Pro Thr Ser  
 1 5 10 15  
 Gln Ala Leu Tyr Ser Asp Phe Ser Pro Pro Glu Gly Leu Glu Glu Leu  
 20 25 30  
 Leu Ser Ala Pro Pro Pro Asp Leu Val Ala Gln Arg His His Gly Trp  
 35 40 45  
 Asn Pro Lys Asp Cys Ser Glu Asn Ile Asp Val Lys Glu Gly Gly Leu  
 50 55 60  
 Cys Phe Glu Arg Arg Pro Val Ala Gln Ser Thr Asp Gly Val Arg Gly  
 65 70 75 80  
 Lys Arg Gly Tyr Ser Arg Gly Leu His Ala Trp Glu Ile Ser Trp Pro  
 85 90 95  
 Leu Glu Gln Arg Gly Thr His Ala Val Val Gly Val Ala Thr Ala Leu  
 100 105 110

261E01-09529660

G1

- 175 -

Ala Pro Leu Gln Ala Asp His Tyr Ala Ala Leu Leu Gly Ser Asn Ser  
 115 120 125

Glu Ser Trp Gly Trp Asp Ile Gly Arg Gly Lys Leu Tyr His Gln Ser  
 130 135 140

Lys Gly Leu Glu Ala Pro Gln Tyr Pro Ala Gly Pro Gln Gly Glu Gln  
 145 150 155 160

Leu Val Val Pro Glu Arg Leu Leu Val Val Leu Asp Met Glu Glu Gly  
 165 170 175

Thr Leu Gly Tyr Ser Ile Gly Gly Thr Tyr Leu Gly Pro Ala Phe Arg  
 180 185 190

Gly Leu Lys Gly Arg Thr Leu Tyr Pro Ser Val Ser Ala Val Trp Gly  
 195 200 205

Gln Cys Gln Val Arg Ile Arg Tyr Met Gly Glu Arg Arg Val Glu Glu  
 210 215 220

Pro Gln Ser Leu Leu His Leu Ser Arg Leu Cys Val Arg His Ala Leu  
 225 230 235 240

Gly Asp Thr Arg Leu Gly Gln Ile Ser Thr Leu Pro Leu Pro Pro Ala  
 245 250 255

Met Lys Arg Tyr Leu Leu Tyr Lys  
 260

## (2) INFORMATION FOR SEQ ID NO:47:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5615 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

00962560-103197

- 176 -

GTACTTTCTT TATATCTCCA TAATTTTATT TACTATTACT ACATGATACA TTATTTTATA	60
AAAGTCTTTG TAACCTCCTT AAGGATTCAC TGCTTAATCT CCAGTGCTTA GCACAAATCA	120
TTAAATGCGA ACCAGAACT CTTCCAAATG TGTTACATCT ATAACCTCAT TGGATTCTCA	180
CTACCAACCC CATGCAATAG ATACTAATGT GATCTCTGTC TTACAGAGGA AGAAACAGGC	240
ACAGGGAGGT TCAGTAATTT GCCCAAGGTC ATACACACAC TGGCCTTCAG GTATTCATGC	300
CCGGGGAGTC TGGTCCCACA GCTGGCATGT TTGCCATTAT ATTATATTGC CTCCTTATAG	360
TGTCGGCACT CATTAAGCAC ATTGACAGCT ATGCTTGGTG AGTGACTACT ATGTACCCAG	420
CTCTGTGCTA CATGCTTTAC CTGGATTATT TCAACTGCAC AACAACCCTG TGAGGTAAC	480
ACCATCATTG CTCCTATTTT ACATAACAGA AACTACAGA AATCTGGGGC TGGGCGTAGT	540
GGCTCATGCC TGAAATCCCA GCACTTTGGG AGACCCTGTC TCTAAAAAA ATTTTTTTTT	600
GGCCGGACGT GGTGGCTCAC ACCTGTAATC TCAGCACTTT GGGAGGCTAA GGCAGGCAGA	660
TCACAAGGTC AGGAGTTCTA GACCAGCCTG GCCAACATGG CAAAACCCTG TGTCTACTAA	720
AAATACAAAA AATAGCTAGG CGTGGTGGCA GGTGCCTGTA ATCCCAGCTA CTCAGGAGGC	780
TGAGGCAGGA GAATCCCCTG AACCTGGGAG ATGGAGGTTA CAGAGAGCCG AGATCGTGCC	840
GCTGCACTCC AGCCTGGGCA ACAAGAGCAA GACTCTGTCT CGAAAAAAT AAAAATAAAA	900
ATAAAAATAT TTTTTTAAAA ATTAGCTGGG TGTGGTAGCA CATGCCTGTA GTCCCAGCTA	960
CTTGGGAGGC TGAGGTAGGA GGATCACTTG AGCCCAGGAG GTCAAGGCTG CAGTGGGCTG	1020
TGATGGCGCC ACTGCACTCT AGCCTTGGTG ACAGCAAGAC CCTGTCTCAA AAAAAAAAAA	1080
AAGAGAAATC GGGCAACTTC CCCAAGATCG CGCAGTTAAC TAGTGGCATA GCTTCACTCA	1140
AACTCGAAGT CTTAATCAGG AACTCTACC AAATGAGATC AACGGCTCAG TAATGGATTG	1200
GCATCCAGTA TGAAGACTGG ACCAGCAGGG AGAACTATGA TGCCTACAGC CTAGAGCCTG	1260
AAGCAGATTT CACAGCCTCA GAGGTGGCAC AGGCTGACTC ACAACCCGGG GCAGAAAGGG	1320
ACCAGCCCAG AAACAGTGAC CCAGAATCAC AGGGAAGTAG AAATGGGATT CGGCACAATG	1380

08962560-103197

- 177 -

AAGCCCCCTCC TTGACCCCAT GCTCCTTACC CTCAGGGGCG CAGGAGTTAG TCGCTCAGGC 1440  
 GGCTCAAAGG TCTTGACGGT GGAGAACACC ATCCCAGGG ATTCCGACG CGGTGATGCC 1500  
 ATCAAAGCGT TAATTCTGAG ATGGGCCTGC CCGGGTGGG ACTCTGCCGC AGCAAGAGAA 1560  
 GGGTTAACTG CCCCGGGCCT TCGCCGTGGG GCGGGGCGCT CGGGGAGGGT CACAGCCCGG 1620  
 GACTGAGACC CGAGGTTAAC CGCCCGGGGT GGGCTCCACG GGGGCGGGG ATGCTCTCCG 1680  
 CGGCTGCTGC CGGTATAGAG CGGTAAGTGC CCAGGAGGGG GCGGGGCCCC ACAGGGGCGT 1740  
 GGCTCGGAG CTGCACGGCC GTGGGCGGCG ATGAGAGGGT TAAGCCCCAG AGGGCCCTGG 1800  
 AGGGGCGGGG CCGCGGGACG GGCTCGGCCC AAGGGAGGAG CTGGGGGCGG AAGCGGCGG 1860  
 CGGTCTGCGC CCTGCGCGCC TCGGCTTCTT TCCGCCCGGC TCCTTCAGAG GCGCGCGAC 1920  
 CTCCAGGGCT GGGAAGTCAA CCGAGGTTCG GGGGCAGCGG CGAGGGCTCC GGGCGAGTAA 1980  
 GGGGGATGGT CCATGCTGAG GCCCAAATGG GGCGAAGTGC CGAGAGTCTC TGGCGACCTG 2040  
 GATCAGATGG GGCAGAGGCA GATGAAGGGC CCAGGAGCTT TGGGGCAGCG AGGAGGGAGG 2100  
 AGCGGGCCCG TTGGCAAAC TGGGTGAAAG GATGGGTAC CTGGGTGACG AGCCCCCGCC 2160  
 AGGATTCTGC TCTTCACGCC CCTTTCTCC CAGCTCCCTT CCAGGTCAAT CCAAAGTGA 2220  
 GCTCAAGTTT CAGAAGAGAA AGACGCCCCA GCAAGCCTCT TCGGGGAGT CCTCTAGCTC 2280  
 CTCACCTCCA TGGGCCAGAC AGCTCTGGCA GGGGGCAGCA GCAGCAGCCC CACGCCACAG 2340  
 GCCCTGTACC CTGACCTCTC CTGTCCCGAG GGCTTGAAG AGCTGCTGTC TGCACCCCT 2400  
 CCTGACCTGG GGGCCAGCG GCGCCACGGT TGGAAACCCA AAGACTGTTC AGAGAACATC 2460  
 GAGGTCAAGG AAGGAGGGTT GTACTTTGAG CGGCGGCCCC TGGCCCAGAG CACTGATGGG 2520  
 GCGGGGGTA AGAGGGGCTA TTCAAGGGGC CTGCACGCCT GGGAGATCAG CTGGCCCCTA 2580  
 GAGCAGAGGG GCACGCATGC CGTGGTGGGC GTGGCCACGG CCCTCGCCCC GCTGCAGACT 2640  
 GACCACTACG CGGCGCTGCT GGGCAGCAAC AGCGAGTCGT GGGGCTGGGA CATCGGGCGG 2700  
 GGGAAGCTGT ACCATCAGAG CAAGGGGCCC GGAGCCCCC AGTATCCAGC GGGAACTCAG 2760

00962560 . 103197

- 178 -

GGTGAGCAGC TGGAGGTGCC AGAGAGACTG CTGGTGGTTC TGGACATGGA GGAGGGAAC 2820

CTGGGCTACG CTATTGGGGG CACCTACCTG GGGCCAGCAT TCCGCGGACT GAAGGGCAGG 2880

ACCCTCTATC CGGCAGTAAG CGCTGTCTGG GGCCAGTGCC AGGTCCGCAT CCGCTACCTG 2940

GGCGAAAGGA GAGGTGAGGC CTGGGGCAGA CGTGGGGAGA ACTTTCGTGC CCTGGTGGCA 3000

GTGGTTTGGG ATGGAACTC TTCTGACAAG AGCAGAGGGG ATGGACCTTC ATCCAGCCTG 3060

CCTCAACCTC TGTTCACTGC TGGGAAAGGC TAGGGGTCTT CACAGCTGTT ATTTAATTTA 3120

ACCCAACAGC AATAGAGGTG AAACAGGCTT GAGAAAGCAA CTTTCTCAAG TTCTCTTGGC 3180

CAGTAAATGG TGAACCTTCA GAATGGAGGG AGGAACTGCA GGGATGAGAG AATTCAGGAG 3240

ATATCAACCC CTGAGCAAGA GGTGCAAAGC GTTAGGTACT GGGTTTGATG TACAGGTCCA 3300

AAAGAAGGAT GGGCAGAGCC AGGTACCCAG GCTGTATACC GGATCCCTG GGCTCTAACC 3360

TGTCTCTGTG CCACATACCT ACTTCCTTCC TCAGCCACAC CTCTGGATGG AGACACTGGG 3420

GCCCTGGGCA CCAGGGAGGA GAGCAGTGGA GGAGGCAGGG CTTAGGGTG GGGCAGCAGG 3480

GGAGGAGCCT CCACAGGAAC TGA CTGGGTC CAGGGCTTGG AGCTGCTCTC TGCAGTTGTG 3540

TGGGCTGTAG AGTGGAGGGC CATCCCTCCT CACCTCAGCC CCAGCTCCCA AGCCTCTGGA 3600

GTCAAAGCCT GGGCCAGCTC CACCACTGTC AGAGCCACCT TGGCCTGTTG TTTAGAGGGC 3660

CTTAGCCAGC TCTTACCCC CAGCTCTGAC TAGGGATGTG TGAAATCTTA TCTGGGAGGC 3720

AGAACTTCCG GGTATCTCAA ATTCCCTTT CAGCCAGGTG GGCACACTCG AAGCAGGAAA 3780

GCAGAAAGGC ATCTGAGTAG GACCCCGTAG TTTGAGGACA TCTGGCTGGT GGCTGCACCC 3840

ATACTTACAT TCCCTCCTT CTCTCTCCA GCGGAGCCAC ACTCCCTTCT GCACCTGAGC 3900

CGCCTGTGTG TGGCCACAA CCTGGGGGAT ACCCGGCTCG GCCAGGTGTC TGCCCTGCCC 3960

TGCCCCCTG CCATGAAGCG CTACCTGCTC TACCAGTGAG CCCTGTGATA CCACAGACTG 4020

TGCTGAGGTC TTGCCACCAC CCCTCCCTT GGGGAGGTGG GGAGGCACTG CTGGCCTAGA 4080

CCAGCTGCTG AAAGCTGGTG AGGCTGAGCC CCTACCCCAA CCCAAGCTCT GCGGAAATCA 4140

08962560-103197



- 179 -

ACAGCCCCAG AGCCACTTGG AGGGAGGAAG AAAGGGAGCC GGCCTTCAAG GCTATGACAG 4200  
 TCTGCTACGC AAAACATTTT TTCAAGTAAA AATAGTAAGA GATGTTGTTA TAGAAACCTG 4260  
 TTCTTGTTTT TTTTITTTTC TTGCACAAAT GATCATTTAT ATAGCTGCCT CAAAAAGGAA 4320  
 GATTATCTGG GCAAGTCCAG TGAAGGCAGA CAAACCACAA GACCTAGTGC CAGGTTTATT 4380  
 CCCTCACATG GGTGTTTAC ATACACAGCA CAGAGGCACG GGCACCATGG GAGAGGGCAG 4440  
 CACTCCTGCC TTCTGAGGGG ATCTTGGCCT CACGGTGTA GAAGGGAGAG GATGGTTTCT 4500  
 CTTCTGCCCT CACTAGGGCC TAGGGAACCC AGGAGCAAAT CCCACCACGC CTTCCATCTC 4560  
 TCAGCCAAGG AGAAGCCACC TTGGTGACGT TTAGTTCCAA CCATTATAGT AAGTGGAGAA 4620  
 GGGATTGGCC TGGTCCCAAC CATTACAGGG TGAAGATATA AACAGTAAAG GAAGATACAG 4680  
 TTTGGATGAG GCCACAGGAA GGAGCAGATG ACACCATCAG AAGCATATGC AGGGAAAGGG 4740  
 CAGTTACTGG GCTTCTGGC TGCTTAGTCC CTGGCTTGGC AGGAAGGGTA GGAAGATGG 4800  
 ATGGGGCTCA TTGTTTGCA TTGATGATGT CCACGAATTC GGGCTTGAGG GAAGCACCAC 4860  
 CCACAAGGAA GCCATCCACA TCAGGCTGGC TGGCCAGCTC CTTGCAGGTT GCCCCAGTCA 4920  
 CAGAGCCTGG GAAGGGAGCA GAACAAGGGC TTGGTCAAGA ATGGGATGAG TCTGCCCCAT 4980  
 CCCCAGCTCC ATGTCCGAGG GCTCAGTCTA GTCCTCAGCC CACTCCACCT CAGCCGGGAA 5040  
 CCAAAGCCAC TCACCTCCAT AAATGATACG GGTGCTCTGA GCCACCGCAT CAGAGACGTT 5100  
 GGACTTCAGC CATCCTCGGA GCTTCTCGTG TACTTCCTGG GCCTAGAACA AGAAGCTGGC 5160  
 CTAAGTAAGA CCTTTTCTGC CTCTCTAAGA GGAAAAATCA CTGGCACCAG TGGACACTTA 5220  
 GTGTGGTTTC TGAAGTATC AGAGTACCAG GGCTCTGATC CAAGCCAGGC CCTGGACTGG 5280  
 ATGCCCTTGG ACAAGTCACT GTCTCTGGGT TCAAGGTCTC TGTGTCTTTG AAATAAGGGG 5340  
 TTGCCCCATG TGGGCTGTGT CTGTCCAAAC CTATTGAGGC AGGCTGGGAT GAGGGCAGGG 5400  
 CTCTTGGGCC CGGTTACCTG TTGGGGTGTT GCAGTCTTGC CAGTACCAAT GGCCACACA 5460  
 GGCTCATAGG CCAGGACGAC CTTGCTCCAG TCCTTCACGT TATCTGCAGG GCAGAGATAC 5520

08962560-103197

- 180 -

AGATGGAGGG AAGGGTGAAC AAGAAAGAGC TCTCCAGCCA GGTTCTCCGG AGTACGAAGA 5580

ACGGTGGCCT ACTGCCCCCT AGTGGACATT GGGGG 5615

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met Gly Gln Thr Ala Leu Ala Gly Gly Ser Ser Ser Thr Pro Thr Pro  
1 5 10 15

Gln Ala Leu Tyr Pro Asp Leu Ser Cys Pro Glu Gly Leu Glu Glu Leu  
20 25 30

Leu Ser Ala Pro Pro Pro Asp Leu Gly Ala Gln Arg Arg His Gly Trp  
35 40 45

Asn Pro Lys Asp Cys Ser Glu Asn Ile Glu Val Lys Glu Gly Gly Leu  
50 55 60

Tyr Phe Glu Arg Arg Pro Val Ala Gln Ser Thr Asp Gly Ala Arg Gly  
65 70 75 80

Lys Arg Gly Tyr Ser Arg Gly Leu His Ala Trp Glu Ile Ser Trp Pro  
85 90 95

Leu Glu Gln Arg Gly Thr His Ala Val Val Gly Val Ala Thr Ala Leu  
100 105 110

Ala Pro Leu Gln Thr Asp His Tyr Ala Ala Leu Leu Gly Ser Asn Ser  
115 120 125

Glu Ser Trp Gly Trp Asp Ile Gly Arg Gly Lys Leu Tyr His Gln Ser

08962560-103197

- 181 -

130 135 140

Lys Gly Pro Gly Ala Pro Gln Tyr Pro Ala Gly Thr Gln Gly Glu Gln  
 145 150 155 160

Leu Glu Val Pro Glu Arg Leu Leu Val Val Leu Asp Met Glu Glu Gly  
 165 170 175

Thr Leu Gly Tyr Ala Ile Gly Gly Thr Tyr Leu Gly Pro Ala Phe Arg  
 180 185 190

Gly Leu Lys Gly Arg Thr Leu Tyr Pro Ala Val Ser Ala Val Trp Gly  
 195 200 205

Gln Cys Gln Val Arg Ile Arg Tyr Leu Gly Glu Arg Arg Ala Glu Pro  
 210 215 220

His Ser Leu Leu His Leu Ser Arg Leu Cys Val Arg His Asn Leu Gly  
 225 230 235 240

Asp Thr Arg Leu Gly Gln Val Ser Ala Leu Pro Leu Pro Pro Ala Met  
 245 250 255

Lys Arg Tyr Leu Leu Tyr Gln  
 260

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGCTAGATCTGGACCCTACAATGGCAGC

28

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs

08962550-103197

- 182 -

- (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGCTAGATCT GCCATCCTAC TCGAGGGGCC AGCTGG

36

03962560-103197